

# Metodología para el Diseño Muestral para la Evaluación de las Declaratorias de Viabilidad otorgadas en el marco del Sistema Nacional de Inversión Pública

Ministerio de Economía y Finanzas  
Dirección General de Programación Multianual del Sector Público

## Índice

Introducción.....	3
I. Objetivos .....	4
II. Determinación del tamaño mínimo de muestra: una breve referencia teórica ...	5
III. Descripción general del sistema de muestreo y unidades de análisis .....	8
IV. Población de trabajo y diseño muestral para la evaluación de declaratorias de viabilidad.....	11
IV.1 Determinación de la población de trabajo .....	11
IV.2 Procedimiento para el trabajo con el sistema y diseño muestral para las declaratorias de viabilidad otorgadas cada año.....	11
IV.2.1 Preparando el sistema con los datos la población de trabajo .....	11
IV.2.2 Procedimiento para el uso de parámetros y niveles de análisis para definir el diseño muestral para la evaluación de declaratorias de viabilidad.....	13
Anexo 1: Clasificación sectorial propuesta para cada función y programa .....	20

## Introducción

1. El Reglamento del Sistema Nacional de Inversión Pública (SNIP) aprobado por Decreto Supremo No. 102-2007 y modificado por el DS 038-2009, en su artículo 3, señala las competencias vigentes de la DGPM. Entre muchas otras, se precisa que dicha Dirección debe realizar evaluaciones muestrales, con periodicidad anual, sobre la calidad de las declaraciones de viabilidad que otorguen los órganos competentes del SNIP.
2. El número de proyectos de inversión pública (PIPs) declarados viables cada año es, sin embargo, demasiado grande como para realizar una evaluación sobre toda la población. De allí que resulta importante establecer una metodología que permita la obtención de una muestra estadísticamente representativa del universo de PIPs declarados viables anualmente.
3. En ese sentido, el presente documento tiene como propósito establecer el procedimiento del diseño muestral que la DGPM debe utilizar cada año para obtener los tamaños de muestra a emplear, de acuerdo a determinados parámetros y niveles de análisis, para que los resultados de la evaluación de declaratorias de viabilidad sean representativos y generalizables para distintos grupos de interés (según nivel de gobierno en el que se sitúa la OPI y dominio geográfico).
4. Con ello se busca hacer transparente el proceso de evaluación de las declaratorias de viabilidades, así como uniformizar criterios técnicos para los posteriores ejercicios de evaluaciones de declaratorias de viabilidades. De este modo, se facilita el análisis comparativo anual del desempeño de las OPIs en su tarea de evaluar los estudios de preinversión que sustentan la declaratoria de viabilidad de los proyectos de inversión pública.

I. Objetivos

5. El objetivo del documento es sustentar un sistema de muestreo que permita a la DGPM obtener los tamaños de muestra a emplear para que los resultados de la evaluación de declaratorias de viabilidad puedan ser generalizables en distintos grupos de interés, de tal modo que se aplique a una población de trabajo debidamente definida.
6. El sistema de muestreo permite hallar el tamaño de muestra a emplear para distintos niveles de análisis: (i) el nivel más agregado corresponde al ámbito de la OPI evaluadora (Gobierno Nacional, Gobierno Regional, Gobierno Local y Empresas). (ii) al interior de cada ámbito, a su vez, es posible evaluar una desagregación en siete sectores priorizados y, (iii) para los ámbitos de gobierno regional y local es posible, también, evaluar una clasificación a nivel de regiones o grupos de regiones.
7. La evaluación a realizar con la muestra nos sugiere dos elementos clave que han servido de guía para el estudio. En primer lugar, es importante notar que la evaluación busca, primordialmente, determinar qué porcentaje de los PIPs declarados viables cumplieron con las condiciones mínimas para serlo. Por lo mismo, el estadístico básico de análisis lo constituye la proporción muestral. Es decir, a partir de la muestra será posible calcular una proporción de PIPs cuya viabilidad fue razonablemente otorgada, y el tamaño de esta muestra deberá ser tal que permita generalizar el resultado obtenido a toda la población, con un margen de error razonable.
8. El segundo elemento a tomar en cuenta es que lo que se evaluará es el desempeño de las OPIs. Por lo mismo, las unidades de análisis al interior de la población de trabajo deberán estar constituidas por los distintos grupos de PIPs que es posible construir a partir de las características de la OPI que declaró su viabilidad. Al respecto, los niveles (ámbitos), sectores y ubicación geográfica de la OPI constituyen los factores discriminantes ideales para distinguir entre distintos grupos de PIPs.

## II. Determinación del tamaño mínimo de muestra: una breve referencia teórica

9. Cuando utilizamos una muestra para calcular determinado estadístico, es importante recordar que el valor que obtenemos es la realización de una variable aleatoria la que, típicamente, responde a una distribución con media y varianza conocidas. En el caso que aquí nos concierne, y tal como fue explicado en el acápite anterior, el principal estadístico de análisis lo constituye la proporción muestral. Este estadístico, denotado comúnmente como  $\hat{p}$ , se distribuye como una Normal con media y varianza conocidas. Formalmente:

$$\hat{p} \sim N\left[p, \frac{p(1-p)}{n}\right] \quad (1.)$$

Donde  $p$  representa la proporción poblacional y  $n$  es el tamaño de muestra utilizado para construir el estadístico. Así, el error (o desviación) estándar de la proporción muestral viene dado por:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \quad (2.)$$

y, en caso la población sea finita ( $N$ ), será necesario agregar un término de corrección de la forma:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}} \quad (3.)$$

10. Ahora bien, preocuparse por el tamaño de muestra a utilizar equivale a preocuparse por la precisión de nuestras estimaciones o, lo que es lo mismo, por el margen de error que estamos dispuestos a tolerar exista entre la estimación y el verdadero valor del parámetro en la población. En nuestro caso estamos hablando de la diferencia entre  $\hat{p}$  y  $p$ .
11. Esta precisión es típicamente medida a través de un intervalo de confianza. En general, dicho intervalo nos indica el rango de valores entre los cuales puedo esperar encontrar el verdadero valor del parámetro para una probabilidad o nivel de confianza dado. Así, para un mismo nivel de confianza, una estimación más precisa será aquella que exhiba un intervalo más acotado.
12. A partir de la distribución muestral de la proporción, es posible construir un intervalo de confianza para dicho estadístico de la forma:

$$\Pr\left[\hat{p} - z_{\alpha/2}\sigma_{\hat{p}} \leq p \leq \hat{p} + z_{\alpha/2}\sigma_{\hat{p}}\right] = (1 - \alpha) \quad (4.)$$

Donde  $z_{\alpha/2}$  se refiere al valor con una probabilidad asociada de  $100(1 - \alpha / 2) \%$  en la distribución normal estándar.

13. La expresión dada en (4.) nos dice que, en un proceso de muestreo repetido, el  $100(1 - \alpha) \%$  de las veces la distancia entre nuestro estimado ( $\hat{p}$ ) y el verdadero valor del parámetro ( $p$ ) no excederá la magnitud  $z_{\alpha/2}\sigma_{\hat{p}}$ . Es, como se dijo, una medida de la precisión de nuestra estimación. Como la magnitud  $z_{\alpha/2}\sigma_{\hat{p}}$  es la que determina la amplitud del intervalo, la precisión dependerá inversamente de ella.
14. Antes de construir la muestra, y al momento de pensar en su tamaño, el investigador busca ejercer control sobre la magnitud  $z_{\alpha/2}\sigma_{\hat{p}}$ , a la que se denomina margen de error ( $e$ ).

$$\begin{aligned} e &= z_{\alpha/2}\sigma_{\hat{p}} \\ &= z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}} \end{aligned} \quad (5.)$$

15. Así, y para una confianza dada (o un valor para  $z_{\alpha/2}$  fijo), calcular el tamaño de muestra a emplear implica determinar el valor de  $n$  que garantiza una desviación estándar ( $\sigma_{\hat{p}}$ ) tal que el margen de error sea de  $100(e)$  puntos porcentuales. En la medida en que se desee un margen de error menor será necesario reducir la desviación estándar y, por lo mismo, incrementar el tamaño de muestra.
16. Tal como se indica en (5.), el margen de error depende de dos parámetros adicionales sobre los que es necesario tomar una decisión:  $z_{\alpha/2}$  y  $p$ . El primero tiene que ver directamente con el nivel de confianza elegido para el intervalo, mientras que el segundo es uno de los argumentos que determinan la desviación estándar del estadístico.
17. Respecto al primero, lo común es elegir un valor para  $\alpha$  de 0.05 y, con esto, una confianza del 95% y un valor para  $z_{\alpha/2}$  de 1.96. En lo que respecta al segundo parámetro, es importante notar que la desviación estándar es creciente en  $p(1-p)$ ; es decir, conforme

mayor sea el valor de  $p(1-p)$ , mayor tendrá que ser el tamaño de muestra para mantener el margen de error constante. Aunado a esto, conviene reparar en que  $p(1-p)$  alcanza su valor máximo (0.25) cuando  $P$  es igual a 0.5. Por lo mismo, y en ausencia de información a priori sobre  $P$  (recordemos que se trata del parámetro que interesa estimar) lo más conservador es utilizar un valor 0.5.

18. Si, no obstante, se sabe que  $P$  se encuentra dentro de un intervalo que no contiene a 0.5, lo recomendable es utilizar el valor límite del intervalo más cercano a 0.5. Por ejemplo, si se sabe que  $0.2 \leq p \leq 0.6$ , se recomienda reemplazar  $p = 0.5$ . No obstante, si se sabe que  $0.1 \leq p \leq 0.3$ , se recomienda trabajar con  $p = 0.3$ .
19. Con estas nociones en mente, basta con despejar (5.) de modo que quede explícito que lo que nos interesa es determinar el valor de  $n$  que garantice el margen de error elegido.

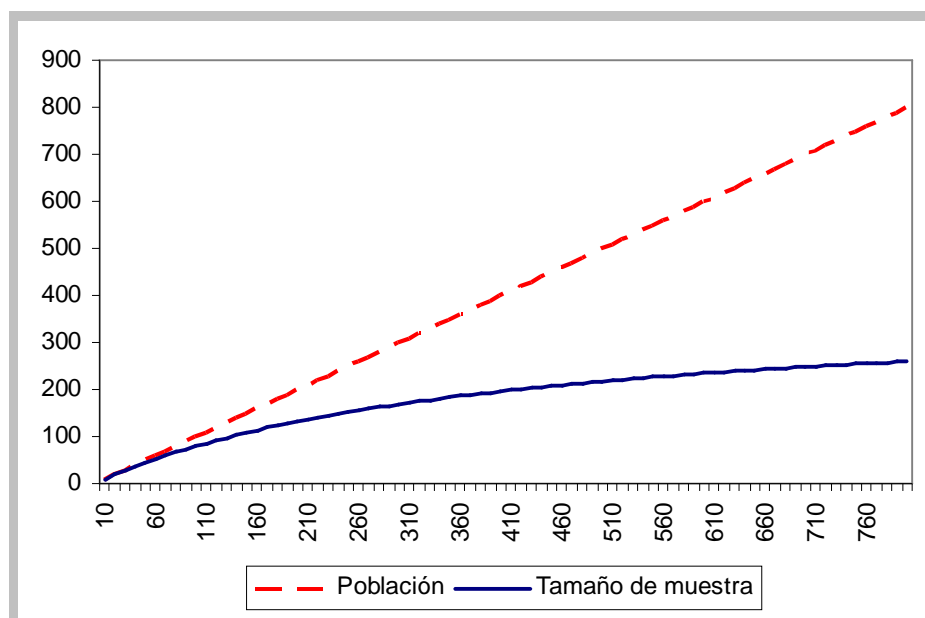
$$\begin{aligned} e &= z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}} \\ e^2 &= (z_{\alpha/2})^2 \left[ \frac{p(1-p)}{n} \right] \left[ \frac{N-n}{N-1} \right] \\ n &= \frac{(z_{\alpha/2})^2 p(1-p)}{e^2} \left[ \frac{N-n}{N-1} \right] \\ n &= n_0 \left[ \frac{N-n}{N-1} \right] \\ n &= \frac{n_0 N}{N-1+n_0} \end{aligned} \tag{6.}$$

Donde  $n_0$  corresponde al tamaño de muestra para una población infinita.

20. Si bien la expresión anterior involucra al tamaño de la población ( $N$ ) esto no implica que el tamaño de muestra a emplear sea una proporción de la misma. Claramente, la relación entre  $n$  y  $N$  no es lineal (ver Gráfico No. 1). Al respecto, conviene pensar en el tamaño de muestra como un concepto independiente del tamaño de la población. En la práctica, el término de corrección por población finita sólo garantizará que la muestra requerida no exceda a la población disponible y no debemos sorprendernos si, para poblaciones pequeñas,  $n$  sea muy cercano a  $N$ . De hecho, lo que este resultado nos está diciendo es

que, dado lo limitado de la información disponible, debemos aprovecharla al máximo para lograr la precisión deseada.

Gráfico No. 1: Tamaño de muestra para distintas poblaciones finitas  
( $\alpha = 5\%$ ;  $p = 0.5$ )



### III. Descripción general del sistema de muestreo y unidades de análisis

21. Tal como se mencionó en la sección introductoria, la DGPM busca evaluar las declaratorias de viabilidad al interior de diversos grupos de interés. Es, por tanto, necesario contar con una herramienta flexible que permita conocer el tamaño de muestra a emplear en cada grupo para distintos niveles de análisis. En esta sección se describen los diversos niveles de análisis y opciones de muestreo disponibles.
22. La unidad básica de análisis propuesta es el ámbito de la OPI evaluadora. Éste puede ser al nivel de Gobierno Nacional, Gobierno Regional, Gobierno Local o Empresas<sup>1</sup>. Esto implica que el número mínimo de proyectos a evaluar provisto por el sistema garantiza que se pueda realizar inferencia (con el margen de error elegido) dentro de estos cuatro grandes grupos.
23. Dentro de cada uno de estos cuatro grupos es posible, a su vez, incrementar el tamaño de muestra para realizar inferencia en los siete sectores priorizados: Agricultura, Educación,

<sup>1</sup> El ámbito “Empresas” agrupa a las OPIs de las empresas que ejecutan PIPs. Básicamente, se trata de empresas proveedoras de energía, agua potable y alcantarillado.



Energía, Institucional, Salud, Saneamiento y Transporte<sup>2</sup>. Esto implica determinar el tamaño mínimo de muestra para cada sector, en cada ámbito de OPI.

24. Los ámbitos de Gobierno Regional y Gobierno Local, por su parte, ofrecen una dimensión de análisis adicional a la sectorial: la dimensión geográfica. Es decir, es posible asociar cada proyecto no sólo a un sector sino también a determinada región geográfica. Esto permite enriquecer el análisis asociado a este tipo de OPIs de dos maneras distintas: utilizando a las regiones como unidades para un muestreo estratificado y/o proponiendo a las regiones (o a grupos de ellas) como unidades de análisis para el proceso de inferencia.
25. Respecto a lo primero, y ya sea que se desee hacer inferencia sólo a nivel de ámbito OPI o se elija una desagregación sectorial para la misma, el sistema procederá a realizar un muestreo estratificado con reparto proporcional según región geográfica. Es decir, las observaciones dentro de cada submuestra (ámbito OPI o ámbito OPI/sector) serán asignadas a cada una de las 25 regiones respetando la participación (o tamaño relativo) de cada región en la población de trabajo. El muestreo estratificado permite incrementar la precisión de nuestros estimados (reducir la varianza del estadístico construido para cada submuestra) pero no permite hacer inferencia al interior de cada estrato o grupo. Por lo mismo, esta opción (activada por defecto) no implica un incremento en el tamaño de muestra.
26. Lo que sí implica incrementar el número de proyectos a evaluar es la posibilidad de realizar inferencia también a nivel geográfico. Para esto, el sistema ofrece dos tipos de unidad geográfica: regiones y grupos. Ésta última representa una opción de trabajo más agregado y requiere que el usuario asigne, previamente, varias regiones a, por lo menos, dos de los cinco grupos disponibles.
27. Elegir cualquiera de estas dos opciones (regiones o grupos) implica solicitar al sistema que calcule el número adicional de observaciones (adicionales a las necesarias para realizar inferencia a nivel de ámbito OPI o ámbito OPI/sector) requerida para que la representatividad de la muestra pueda ser extendida al nivel geográfico.
28. Los Cuadros No. 1 y No. 2 muestran las unidades de análisis y tipos de muestreo disponibles para cada uno de los cuatro ámbitos OPI considerados. La determinación del número total de proyectos a evaluar pasa por elegir un tipo de desagregación para los ámbitos Gobierno Nacional y Empresas, y un tipo de desagregación y unidad geográfica

---

<sup>2</sup> En el Anexo 1 se muestra la clasificación sectorial para cada función y programa.

para los ámbitos Gobierno Regional y Gobierno Local. Evidentemente, el tamaño total de muestra será mayor conforme el número de unidades de análisis considerado crezca.

Cuadro No. 1: Unidades de análisis para las que es posible realizar inferencia (Ámbito OPI = Gobierno Nacional o Empresas)

Desagregación	Ninguna	► Ámbito OPI (1) con muestreo simple.
	Sectorial	► Ámbito OPI (1) con muestreo simple. ► Sectores (7) con muestreo simple.

Cuadro No. 2: Unidades de análisis para las que es posible realizar inferencia (Ámbito OPI = Gobierno Regional o Gobierno Local)

		Unidad Geográfica		
		Ninguna	Regiones	Grupos
Desagregación	Ninguna	► Ámbito OPI (1) con reparto proporcional por regiones.	► Ámbito OPI (1) con reparto proporcional por regiones. ► Regiones (25) con muestreo simple.	► Ámbito OPI (1) con reparto proporcional por grupos. ► Grupos (hasta 5) con muestreo simple.
	Sectorial	► Ámbito OPI (1) con reparto proporcional por regiones. ► Sectores (7) con reparto proporcional por regiones.	► Ámbito OPI (1) con reparto proporcional por regiones. ► Sectores (7) con reparto proporcional por regiones. ► Regiones (25) con muestreo simple.	► Ámbito OPI (1) con reparto proporcional por grupos. ► Sectores (7) con reparto proporcional por grupos. ► Grupos (hasta 5) con muestreo simple.

IV. Población de trabajo y diseño muestral para la evaluación de declaratorias de viabilidad.

IV.1 Determinación de la población de trabajo

29. La población de trabajo se determinará sobre la base del total de proyectos declarados viables por las OPIs durante el año que se fije como periodo de la evaluación de declaratorias de viabilidad.

30. El propósito de este ejercicio es conocer cuál es el subconjunto de proyectos que, de acuerdo a determinadas características, concentran un monto importante de inversión del presupuesto total de proyectos declarados viables para el año sujeto a análisis, de tal modo de ganar eficiencia acotando el universo de proyectos sobre el cual se obtendrán los tamaños de muestra, con solo una ligera pérdida de representatividad presupuestal.

31. Criterios como el umbral del monto de inversión y la participación de un sector dentro del presupuesto de inversión total declarado viable en el año bajo análisis, podrán ser tomados en cuenta para acotar el universo de proyectos sobre los cuales se obtendrán los tamaños de muestra. Otros criterios podrán ser considerados previo sustento.

IV.2 Procedimiento para el trabajo con el sistema y diseño muestral para las declaratorias de viabilidad otorgadas cada año.

32. El sistema de muestreo ha sido desarrollado en MS Excel e involucra seis hojas con las que debe interactuar el usuario. Tres de ellas son hojas de "insumo" (donde el usuario debe especificar los datos relacionados con la población de trabajo); una de ellas contiene el panel de control (parámetros y niveles de análisis) y un resumen de resultados (el tamaño final de muestra según ámbito OPI y sector); y las dos restantes muestran el detalle de resultados a nivel de regiones y/o grupos.

IV.2.1 Preparando el sistema con los datos la población de trabajo

33. En la hoja (I) N x OPI, el usuario debe ingresar la población de trabajo clasificada según ámbito OPI y sector. En la hoja (I) N x Reg, por su parte, se debe ingresar la población de trabajo asociada a los ámbitos de Gobierno Regional y Gobierno Local, clasificada según región y sector. La imagen a continuación se visualiza una muestra de un extracto de los

datos asociados a una población de trabajo tomado como ejemplo para un determinado año, elegida de acuerdo con los criterios descritos en la sección anterior (punto 4.1). Según el ejemplo indicado en la Imagen No. 1, la población de trabajo de referencia para el año bajo análisis involucra un total de 10,212 proyectos.

Imagen No. 1: Extracto de las hojas para el ingreso de información sobre la población de trabajo

...> Hoja (I) N x OPI

**Población (N) según ámbito de la OPI y sector**

Sector	Ámbito OPI				TOTAL
	Gobierno Nacional	Gobierno Regional	Gobierno Local	Empresas	
1 Agricultura	92	329	913	0	1,334
2 Educación	206	559	1,881	0	2,646
3 Energía	56	110	364	47	577
4 Institucional	15	31	223	3	272
5 Salud	38	121	335	0	494
6 Saneamiento	32	150	1,910	133	2,225
7 Transporte	209	220	2,232	3	2,664
<b>TOTAL</b>	<b>648</b>	<b>1,520</b>	<b>7,858</b>	<b>186</b>	<b>10,212</b>

...> Hoja (I) N x Reg

Ingrese los datos de la población de trabajo aquí

	Gobierno Regional							TOTAL	Gobierno Local
	Agricultura	Educación	Energía	Institucional	Salud	Saneamiento	Transporte		
1 Amazonas	5	32	8	2	6	10	8	71	30
2 Ancash	38	59	3	1	12	2	4	119	150
3 Apurímac	17	8	0	2	5	12	4	48	56
4 Arequipa	30	33	4	2	6	24	21	120	2
5 Ayacucho	46	21	4	2	3	6	16	98	95
6 Cajamarca	17	33	6	0	5	3	14	78	60
7 Callao	0	1	0	0	1	2	6	10	0
8 Cusco	31	57	7	2	7	8	6	118	99
9 Huancavelica	23	53	25	1	4	9	41	156	66
10 Huánuco	9	9	1	0	3	1	3	26	21
11 Ica	6	9	0	0	10	7	5	37	0
12 Junín	16	42	2	1	5	6	10	82	34
13 La Libertad	5	15	4	1	10	4	5	44	46
14 Lambayeque	7	17	1	0	7	1	2	35	7
15 Lima	12	19	6	1	1	10	25	74	20
16 Loreto	2	18	11	1	6	8	3	49	13
17 Madre de Dios	1	5	0	0	1	4	11	22	0
18 Moquegua	5	19	0	1	5	0	6	36	45
19 Pasco	6	6	1	3	0	2	0	18	12
20 Piura	9	32	7	1	11	12	4	76	20
21 Puno	2	12	0	1	2	0	2	19	105
22 San Martín	14	13	15	3	2	12	7	66	11
23 Tacna	6	23	0	1	8	3	4	45	16
24 Tumbes	10	11	3	1	0	2	6	33	2
25 Ucayali	12	12	2	4	1	2	7	40	3
<b>TOTAL</b>	<b>329</b>	<b>559</b>	<b>110</b>	<b>31</b>	<b>121</b>	<b>150</b>	<b>220</b>	<b>1,520</b>	<b>913</b>

34. Si se planea elegir la opción de “Grupos” para la unidad geográfica, en la hoja Regiones&Grupos el usuario deberá especificar a qué grupo pertenece cada región. El sistema admite hasta cinco grupos y una clasificación distinta para OPIs de ámbito regional y local. De acuerdo con lo mostrado en la Imagen No. 2, se propone una misma clasificación para los dos ámbitos que corresponde a los grupos de necesidad establecidos por el Fondo de Promoción a la Inversión Pública Regional y Local (FONIPREL).

Imagen No. 2: Extracto de la hoja Regiones&Grupos para la clasificación de regiones (opcional)

	Gobierno Regional		Gobierno Local	
1 Amazonas	G1	Muy Alta Nec.	G1	Muy Alta Nec.
2 Ancash	G2	Alta Nec.	G2	Alta Nec.
3 Apurímac	G1	Muy Alta Nec.	G1	Muy Alta Nec.
4 Arequipa	G3	Nec. Media	G3	Nec. Media
5 Ayacucho	G1	Muy Alta Nec.	G1	Muy Alta Nec.
6 Cajamarca	G1	Muy Alta Nec.	G1	Muy Alta Nec.
7 Callao	G3	Nec. Media	G3	Nec. Media
8 Cusco	G2	Alta Nec.	G2	Alta Nec.
9 Huancaavelica	G1	Muy Alta Nec.	G1	Muy Alta Nec.
10 Huánuco	G1	Muy Alta Nec.	G1	Muy Alta Nec.
11 Ica	G3	Nec. Media	G3	Nec. Media
12 Junín	G2	Alta Nec.	G2	Alta Nec.
13 La Libertad	G2	Alta Nec.	G2	Alta Nec.
14 Lambayeque	G2	Alta Nec.	G2	Alta Nec.
15 Lima	G3	Nec. Media	G3	Nec. Media
16 Loreto	G1	Muy Alta Nec.	G1	Muy Alta Nec.
17 Madre de Dios	G2	Alta Nec.	G2	Alta Nec.
18 Moquegua	G3	Nec. Media	G3	Nec. Media
19 Pasco	G1	Muy Alta Nec.	G1	Muy Alta Nec.
20 Piura	G2	Alta Nec.	G2	Alta Nec.
21 Puno	G1	Muy Alta Nec.	G1	Muy Alta Nec.
22 San Martín	G2	Alta Nec.	G2	Alta Nec.
23 Tacna	G3	Nec. Media	G3	Nec. Media
24 Tumbes	G3	Nec. Media	G3	Nec. Media
25 Ucayali	G2	Alta Nec.	G2	Alta Nec.

Grupos	Descripción
G1	Muy Alta Nec.
G2	Alta Nec.
G3	Nec. Media
G4	
G5	

Ingrese el nombre de los grupos aquí

Elija a qué grupo pertenece cada región

#### IV.2.2 Procedimiento para el uso de parámetros y niveles de análisis para definir el diseño muestral para la evaluación de declaratorias de viabilidad.

35. Una vez introducida la información sobre la población de trabajo, es necesario elegir los parámetros para el tamaño de muestra así como el nivel de desagregación y unidad geográfica para los distintos ámbitos de OPI (ver Cuadros No. 1 y No. 2).
36. Para fines ilustrativos, se muestra la aplicación de un conjunto de parámetros estándar (margen de error de 5%, nivel de confianza del 95% y proporción poblacional igual a 0.5) y sólo a nivel de ámbito OPI. La imagen siguiente muestra un extracto del panel de control y

del resumen de resultados consistente con los parámetros elegidos (hoja Control&Resumen).

Imagen No. 3: Extracto del panel de control y resumen de resultados (parámetros estándar y nivel de análisis agregado)

The image shows a software interface for the SNIP (Sistema Nacional de Información de Proyectos). It includes a logo, a table for disaggregation by geographic unit, a table for sampling parameters, and a large table for the results summary by sector and geographic scope.

**Desagregación Unidad Geo.**

Ámbito OPI	Desagregación	Unidad Geo.
Gobierno Nacional	Ninguna	NA
Gobierno Regional	Ninguna	Ninguna
Gobierno Local	Ninguna	Ninguna
Empresas	Ninguna	NA

**Parámetros**

Margen de error (e)	5.0%
Nivel de confianza (1- $\alpha$ )	95.0%
Proporción pob. (p)	0.5


**Ámbito OPI**

Sector	Gobierno Nacional	Gobierno Regional	Gobierno Local	Empresas	TOTAL
1 Agricultura	NA	NA	NA	NA	0
2 Educación	NA	NA	NA	NA	0
3 Energía	NA	NA	NA	NA	0
4 Institucional	NA	NA	NA	NA	0
5 Salud	NA	NA	NA	NA	0
6 Saneamiento	NA	NA	NA	NA	0
7 Transporte	NA	NA	NA	NA	0
<b>Total</b>	<b>242</b>	<b>307</b>	<b>368</b>	<b>126</b>	<b>1,043</b>
OverS	NA	0	0	NA	0
<b>TOTAL</b>	<b>242</b>	<b>307</b>	<b>368</b>	<b>126</b>	<b>1,043</b>

37. De acuerdo con los resultados mostrados en la imagen anterior, y de acuerdo a la población de trabajo tomada como ejemplo, para hacer inferencia con un margen de error de 5% en cada ámbito OPI se necesita evaluar un total de 1,043 proyectos: 242 para poder llegar a conclusiones generalizables sobre las OPIs del ámbito de Gobierno Nacional, 307 para el ámbito de Gobierno Regional, 368 para el local y 126 para la categoría Empresas. Según se indica en los Cuadros No. 1 y No. 2, las muestras para los ámbitos regional y local se encuentran, por defecto, estratificadas con un reparto proporcional por regiones. Para conocer el reparto propuesto es necesario acceder a la hoja UGeoReg.
38. Se puede completar la exploración del sistema de muestreo eligiendo parámetros y niveles de análisis más adecuados para la evaluación de los PIPs declarados viables para un año determinado.

39. Se recomienda mantener el nivel de confianza en un 95%. De hecho, no es conveniente alterar este parámetro tomando en cuenta que más adelante se podría modificar el margen de error. Si se está dispuesto a trabajar con una precisión menor, conviene traducir esto directamente en términos de un mayor margen de error para un mismo nivel de confianza. Así, los resultados serán mucho más fáciles de interpretar y, eventualmente, comunicar a terceros.
40. En lo que respecta a la proporción poblacional, se puede emplear la información sobre la estimación de tal proporción de un periodo anterior al año bajo análisis (el cual puede encontrarse dentro de algún rango que no contenga a 0.5). Para fines explicativos, se asignará un valor de 0.2 para la proporción poblacional, asumiendo que el límite superior del rango de resultados de la proporción de proyectos conformes y no conformes del periodo anterior fue de 20%. Esto permitirá reducir el tamaño de muestra requerido manteniendo un enfoque conservador.
41. Debido al interés de conocer la calidad de los estudios de preinversión tanto a nivel sectorial como dentro de unidades geográficas del nivel regional, es posible elegir en el sistema de muestreo estos niveles de análisis. Por lo mismo, se seleccionará una desagregación sectorial para todos los ámbitos OPI. Además, para los de gobierno regional y local se elegirán regiones y grupos como unidad geográfica, respectivamente. Esto último permitirá añadir una dimensión geográfica incluso al nivel local, sin incrementar excesivamente el tamaño de muestra (recordemos que el ámbito local es el que concentra la mayoría de proyectos).
42. Las opciones elegidas en el ejemplo descrito en el punto 35 permitían realizar inferencia en cuatro unidades de análisis: los cuatro distintos ámbitos OPI. Las indicadas en el párrafo anterior (punto 41), en cambio, permiten hacer inferencia en un total de 60 unidades de análisis: (i) el ámbito OPI de Gobierno Nacional y los siete sectores al interior de este ámbito (8 unidades); (ii) el ámbito OPI de Gobierno Regional, los siete sectores al interior de este ámbito y las 25 regiones (33 unidades); (iii) el ámbito OPI de Gobierno Local, los siete sectores al interior de este ámbito y los tres grupos de necesidades FONIPREL (11 unidades); y (iv) el ámbito OPI de Empresas y los siete sectores al interior de este ámbito (8).
43. Como resultado del incremento en el número de unidades de análisis y con la población de trabajo de referencia, el tamaño total de la muestra ha crecido significativamente (de 1,043 a 2,992 proyectos), tal como se muestra en la imagen a continuación:

Imagen No. 5: Extracto del panel de control y resumen de resultados



		Desagregación	Unidad Geo.
Ámbito OPI	Gobierno Nacional	Sectorial	NA
	Gobierno Regional	Sectorial	Regiones
	Gobierno Local	Sectorial	Grupos
	Empresas	Sectorial	NA

Parámetros	
Margen de error (e)	5.0%
Nivel de confianza (1- $\alpha$ )	95.0%
Proporción pob. (p)	0.2

Sector	Ámbito OPI				TOTAL
	Gobierno Nacional	Gobierno Regional	Gobierno Local	Empresas	
1 Agricultura	68	141	194	0	403
2 Educación	113	172	218	0	503
3 Energía	46	79	147	40	312
4 Institucional	15	31	118	3	167
5 Salud	34	81	143	0	258
6 Saneamiento	29	96	218	87	430
7 Transporte	114	117	222	3	456
<b>Total</b>	<b>419</b>	<b>717</b>	<b>1,260</b>	<b>133</b>	<b>2,529</b>
OverS	NA	450	13	NA	463
<b>TOTAL</b>	<b>419</b>	<b>1,167</b>	<b>1,273</b>	<b>133</b>	<b>2,992</b>

44. De acuerdo con la información mostrada en la Imagen No. 5, y con respecto al análisis agregado del acápite anterior, el número de proyectos al interior de cada ámbito OPI se ha incrementado de modo que los resultados sean también generalizables a nivel de cada sector. Así, por ejemplo, ahora se necesita un total de 717 evaluaciones para garantizar que, al interior del ámbito regional, éstas sean representativas en los siete sectores. El cuadro resumen incluido en la hoja Control&Resumen nos indica cómo están distribuidas estas 717 observaciones en los siete sectores. La hoja UGeoReg, por su parte, nos muestra el resultado de repartir las observaciones de cada sector de manera proporcional según las 25 regiones.
45. Para este mismo ámbito, también se ha elegido "Regiones" como unidad geográfica. Al respecto, el cuadro resumen mostrado en la imagen anterior nos indica que para lograr la representatividad regional son necesarias 450 observaciones adicionales para un total de



1,167 proyectos. En la hoja UGeoReg se muestra como están distribuidas estas 450 observaciones adicionales<sup>3</sup>.

Imagen No. 6: Extracto de la hoja UGeoReg, reparto proporcional y observaciones adicionales por región

	Gobierno Regional							Total	OverS	TOTAL
	Agricultura	Educación	Energía	Institucional	Salud	Saneamiento	Transporte			
1 Amazonas	2	10	6	2	4	6	4	34	22	56
2 Ancash	16	18	2	1	8	1	2	48	33	81
3 Apurímac	7	2	0	2	3	8	2	24	17	41
4 Arequipa	13	10	3	2	4	15	11	58	23	81
5 Ayacucho	20	6	3	2	2	4	9	46	25	71
6 Cajamarca	7	10	4	0	3	2	7	33	27	60
7 Callao	0	0	0	0	1	1	3	5	5	10
8 Cusco	13	17	5	2	5	5	3	50	30	80
9 Huancavelica	10	16	18	1	3	6	22	76	20	96
10 Huánuco	4	3	1	0	2	1	2	13	11	24
11 Ica	3	3	0	0	7	4	3	20	13	33
12 Junín	7	13	1	1	3	4	5	34	28	62
13 La Libertad	2	5	3	1	7	3	3	24	14	38
14 Lambayeque	3	5	1	0	5	1	1	16	15	31
15 Lima	5	6	4	1	1	6	13	36	22	58
16 Loreto	1	6	8	1	4	5	2	27	14	41
17 Madre de Dios	0	2	0	0	1	3	6	12	9	21
18 Moquegua	2	6	0	1	3	0	3	15	17	32
19 Pasco	3	2	1	3	0	1	0	10	7	17
20 Piura	4	10	5	1	7	8	2	37	22	59
21 Puno	1	4	0	1	1	0	1	8	10	18
22 San Martín	6	4	11	3	1	8	4	37	16	53
23 Tacna	3	7	0	1	5	2	2	20	19	39
24 Tumbes	4	3	2	1	0	1	3	14	16	30
25 Ucayali	5	4	1	4	1	1	4	20	15	35
<b>TOTAL</b>	<b>141</b>	<b>172</b>	<b>79</b>	<b>31</b>	<b>81</b>	<b>96</b>	<b>117</b>	<b>717</b>	<b>450</b>	<b>1,167</b>

Muestreo estratificado: cómo repartir las observaciones de cada sector según regiones

Representatividad regional: cómo están distribuidas las observaciones adicionales

46. Continuando con nuestro ejemplo, y de acuerdo con lo indicado en el Cuadro No. 2, lo anterior garantiza que los resultados obtenidos a nivel de sector y región sean representativos de sus respectivas poblaciones, con el margen de error elegido (5%).

47. Si es posible realizar inferencia a nivel de sector y región, con mayor razón deberá serlo a nivel de ámbito OPI. De hecho, cabe esperar que el margen de error se haya reducido en la medida en que el número total de observaciones obtenidas de cada ámbito se ha incrementado. El margen de error implícito asociado a cada ámbito agregado se muestra en la tabla de parámetros. Al respecto, cabe precisar que: (i) este margen de error implícito coincidirá con el elegido si no se selecciona ningún tipo de desagregación ni unidad geográfica; (ii) si se selecciona una desagregación sectorial y/o alguna unidad

<sup>3</sup> Esta misma lógica aplica al ámbito OPI de Gobierno Local pero para un análisis geográfico a nivel de grupos (será necesario acceder a la hoja UGeoGr en lugar de UGeoReg).

geográfica, el margen de error implícito a nivel de ámbito OPI será menor al margen de error elegido, y este último aplicará al análisis a nivel sectorial y/o unidad geográfica.

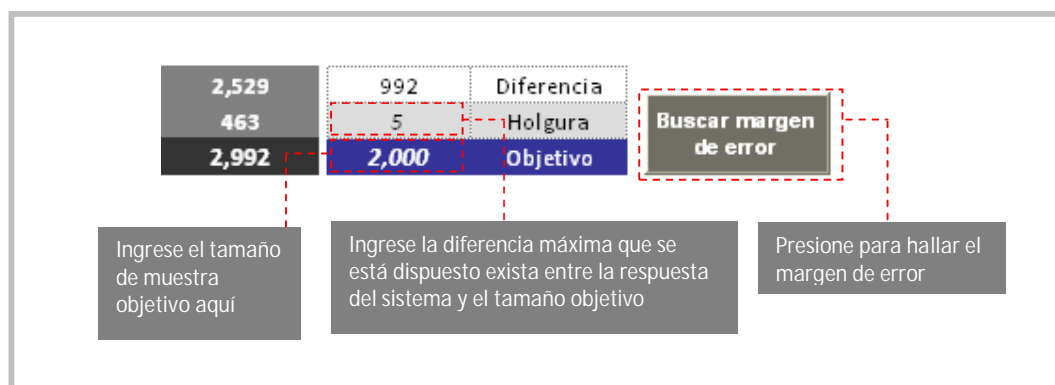
Imagen No. 7: Tabla de parámetros y margen de error implícito a nivel de ámbito OPI

Parámetros		Gob.Nacional	Gob.Regional	Gob.Local	Empresas
Margen de error (e)	5.0%	2.3%	1.1%	2.0%	3.6%
Nivel de confianza (1- $\alpha$ )	95.0%				
Proporción pob. (p)	0.2				

48. Para clarificar lo anterior, retomemos nuestro ejemplo asociado al ámbito de Gobierno Regional y recordemos que el hecho de haber elegido una desagregación sectorial y una unidad geográfica regional resultó en un incremento en el tamaño de muestra (de 307 a 1,167 observaciones para el caso del ejemplo desarrollado). Al respecto, y tal como se muestra en la imagen anterior, este incremento ha permitido reducir el margen de error para el análisis agregado a nivel de ámbito a sólo 1.1%. En contraste, el margen de error elegido (5%) corresponde al análisis a nivel de sector y región.
49. En la columna TOTAL mostrada en la Imagen No. 6 se indica cómo están repartidas las 1,167 observaciones del ámbito OPI Gobierno Regional entre las 25 regiones. Esta distribución es muy similar a la que se obtendría con un reparto proporcional<sup>4</sup>.
50. Por último, cabe destacar un atributo adicional del sistema: la posibilidad de calcular el margen de error que habría que admitir si se quiere que el número total de proyectos a evaluar sea igual a un tamaño de muestra objetivo. Para esto, basta con especificar el tamaño objetivo y la holgura o diferencia máxima que se está dispuesto exista entre la respuesta del sistema y el objetivo especificado.

<sup>4</sup> Nótese que lo que efectivamente corresponde a un reparto proporcional es la distribución al interior de cada sector. Si bien es muy similar, el resultado total no corresponde exactamente a este tipo de reparto.

Imagen No. 8: Opciones para trabajar con un tamaño de muestra objetivo



51. En la imagen anterior se muestra cómo preguntar al sistema cuál sería el margen de error asociado a un tamaño de muestra objetivo de 2,000 evaluaciones, manteniendo las mismas unidades de análisis indicadas líneas arriba. A diferencia del total asociado a un margen de error de 5%, este tamaño objetivo implica sólo un incremento de 33%.
52. La respuesta del sistema se muestra a continuación: para reducir el tamaño de muestra a 2,000 evaluaciones, es necesario admitir un margen de error de 7.4% a nivel de sector, región y grupo. Cabe destacar, sin embargo, que los márgenes de error implícitos a nivel de ámbito OPI se encuentran, en su mayoría, aún por debajo del 5%.

Imagen No. 9: Resultados asociados a un tamaño de muestra objetivo de 2,000 observaciones

Parámetros		Gob.Nacional	Gob.Regional	Gob.Local	Empresas
Margen de error (e)	7.4%	3.3%	1.6%	2.9%	5.2%
Nivel de confianza (1-α)	95.0%				
Proporción pob. (p)	0.2				

Sector	Ámbito OPI				TOTAL
	Gobierno Nacional	Gobierno Regional	Gobierno Local	Empresas	
1 Agricultura	51	86	100	0	237
2 Educación	73	96	105	0	274
3 Energía	38	60	86	34	218
4 Institucional	14	28	75	3	120
5 Salud	29	54	85	0	168
6 Saneamiento	26	63	106	62	257
7 Transporte	74	74	107	3	258
Total	305	461	664	102	1,532
OverS	NA	462	2	NA	464
TOTAL	305	923	666	102	1,996

0	Diferencia
5	Holgura
2,000	Objetivo

Buscar margen de error

## Anexo 1: Clasificación sectorial propuesta para cada función y programa

Función	Programa	Sector
LEGISLATIVA	PROCESO LEGISLATIVO	Otros
JUSTICIA	JUSTICIA	Otros
ADMINISTRACION Y PLANEAMIENTO	ADMINISTRACION	Institucional
ADMINISTRACION Y PLANEAMIENTO	LUCHA CONTRA EL CONSUMO DE DROGAS	Otros
ADMINISTRACION Y PLANEAMIENTO	ADMINISTRACION FINANCIERA	Institucional
ADMINISTRACION Y PLANEAMIENTO	FISCALIZACION FINANCIERA Y PRESUPUESTARIA	Institucional
ADMINISTRACION Y PLANEAMIENTO	PLANEAMIENTO GUBERNAMENTAL	Institucional
ADMINISTRACION Y PLANEAMIENTO	CIENCIA Y TECNOLOGIA	Otros
AGRARIA	ORGANIZACION AGRARIA	Agricultura
AGRARIA	PROMOCION DE LA PRODUCCION AGRARIA	Agricultura
AGRARIA	PROMOCION DE LA PRODUCCION PECUARIA	Agricultura
AGRARIA	PRESERVACION DE LOS RECURSOS NATURALES RENOVABLES	Agricultura
AGRARIA	PROMOCION Y EXTENSION RURAL	Agricultura
ASISTENCIA Y PREVISION SOCIAL	ASISTENCIA SOLIDARIA	Otros
ASISTENCIA Y PREVISION SOCIAL	PROMOCION Y ASISTENCIA SOCIAL COMUNITARIA	Otros
ASISTENCIA Y PREVISION SOCIAL	PREVISION	Otros
COMUNICACIONES	COMUNICACIONES POSTALES	Otros
COMUNICACIONES	TELECOMUNICACIONES	Otros
DEFENSA Y SEGURIDAD NACIONAL	SERVICIOS DE INTELIGENCIA	Otros
DEFENSA Y SEGURIDAD NACIONAL	ORDEN INTERNO	Otros
DEFENSA Y SEGURIDAD NACIONAL	ORDEN EXTERNO	Otros
DEFENSA Y SEGURIDAD NACIONAL	PREVENCION Y ATENCION DE DESASTRES	Otros
DEFENSA Y SEGURIDAD NACIONAL	DEFENSA CONTRA SINIESTROS	Otros
EDUCACION Y CULTURA	EDUCACION INICIAL	Educación
EDUCACION Y CULTURA	EDUCACION PRIMARIA	Educación
EDUCACION Y CULTURA	EDUCACION SECUNDARIA	Educación
EDUCACION Y CULTURA	EDUCACION SUPERIOR	Educación
EDUCACION Y CULTURA	CAPACITACION Y PERFECCIONAMIENTO	Educación
EDUCACION Y CULTURA	EDUCACION ESPECIAL	Educación
EDUCACION Y CULTURA	ASISTENCIA A EDUCANDOS	Educación
EDUCACION Y CULTURA	EDUCACION FISICA Y DEPORTES	Educación
EDUCACION Y CULTURA	CULTURA	Educación

EDUCACION Y CULTURA	INFRAESTRUCTURA EDUCATIVA	Educación
ENERGIA Y RECURSOS MINERALES	ENERGIA	Energía
ENERGIA Y RECURSOS MINERALES	HIDROCARBUROS	Energía
ENERGIA Y RECURSOS MINERALES	RECURSOS MINERALES	Energía
ENERGIA Y RECURSOS MINERALES	RECURSOS HIDRICOS	Energía
INDUSTRIA, COMERCIO Y SERVICIOS	INDUSTRIA	Otros
INDUSTRIA, COMERCIO Y SERVICIOS	COMERCIO	Otros
INDUSTRIA, COMERCIO Y SERVICIOS	TURISMO	Otros
INDUSTRIA, COMERCIO Y SERVICIOS	PROTECCION DE LA LIBRE COMPETENCIA	Otros
PESCA	PROMOCION DE LA PRODUCCION PESQUERA	Otros
RELACIONES EXTERIORES	POLITICA EXTERIOR	Otros
SALUD Y SANEAMIENTO	SANEAMIENTO	Saneamiento y Vivienda
SALUD Y SANEAMIENTO	PROTECCION DEL MEDIO AMBIENTE	Saneamiento y Vivienda
SALUD Y SANEAMIENTO	SALUD COLECTIVA	Salud
<b>Función</b>	<b>Programa</b>	<b>Sector</b>
SALUD Y SANEAMIENTO	SALUD INDIVIDUAL	Salud
TRABAJO	PRESTACIONES LABORALES	Otros
TRABAJO	PROTECCION AL TRABAJADOR	Otros
TRANSPORTE	TRANSPORTE AEREO	Transporte
TRANSPORTE	TRANSPORTE TERRESTRE	Transporte
TRANSPORTE	TRANSPORTE FERROVIARIO	Transporte
TRANSPORTE	TRANSPORTE HIDROVIARIO	Transporte
TRANSPORTE	TRANSPORTE METROPOLITANO	Transporte
VIVIENDA Y DESARROLLO URBANO	VIVIENDA	Saneamiento y Vivienda
VIVIENDA Y DESARROLLO URBANO	DESARROLLO URBANO	Saneamiento y Vivienda
***	***	Blanco