

Departamento Administrativo
Nacional de Estadística



Diseño
DSO

Dirección de Metodología y Producción
Estadística – DIMPE

**Diseño Estadístico
Encuesta sobre Ambiente y Desempeño
Institucional Departamental
-EDID-**

Agosto 2013



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 2
FECHA: 29-08-2013

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

TABLA DE CONTENIDO

1. MARCO ESTADÍSTICO.....	3
1.1 OBJETIVOS.....	3
1.2 TIPO DE OPERACIÓN ESTADÍSTICA.....	3
1.3 UNIVERSO.....	3
1.4 POBLACIÓN OBJETIVO.....	3
1.5. VARIABLES DE INTERÉS.....	3
1.6. PARÁMETROS A ESTIMAR.....	4
1.7. COBERTURA.....	4
1.8. NIVELES DE DESAGREGACIÓN.....	4
1.9. UNIDADES ESTADÍSTICAS.....	4
1.10 PERIODO DE REFERENCIA.....	5
1.11. PERIODO DE RECOLECCIÓN.....	5
2. DISEÑO MUESTRAL.....	6
2.1 MARCO MUESTRAL.....	6
2.2 TIPO DE MUESTREO.....	6
2.3 DEFINICIÓN DEL TAMAÑO DE LA MUESTRA.....	7
2.4 PROCEDIMIENTO DE ESTIMACIÓN.....	8
3. DISEÑO DE INDICADORES.....	14
3.1 INDICADORES SIMPLES.....	14
3.2 INDICADORES SINTÉTICOS.....	14
3.3 CÁLCULO DE INDICADORES SIMPLES.....	14
3.4 CÁLCULO DE INDICADORES SINTÉTICOS.....	15
3.5 TIPOS DE VARIABLES.....	17
3.6 CONFORMACION DE LOS INDICADORES.....	17
BIBLIOGRAFÍA.....	19
ANEXO A. EDID 2013. TAMAÑO DE MUESTRA POR ENTIDAD.....	20

PROCESO: Diseño

SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID

ELABORÓ: Diseños Muestrales

REVISÓ: Coordinación Temática Social

APROBÓ: Director Técnico DIMPE

1. MARCO ESTADISTICO

1.1 OBJETIVOS

Objetivo general

Conocer la percepción de los funcionarios territoriales sobre el ambiente y el desempeño institucional de las 32 Gobernaciones y el Distrito Capital, a las que prestan sus servicios.

Objetivos específicos

- Obtener información sobre la percepción de los funcionarios territoriales respecto al ambiente institucional de las gobernaciones y el Distrito Capital, a partir del conocimiento sobre el nivel existente de credibilidad en las reglas, en las políticas y frente a la suficiencia de los recursos.
- Recoger información sobre la percepción de los funcionarios territoriales respecto al desempeño institucional de las entidades, a través del conocimiento sobre los logros alcanzados en gestión por resultados, rendición de cuentas, bienestar laboral y prácticas irregulares.
- Recoger información sobre la percepción de los funcionarios de las gobernaciones y el Distrito Capital respecto a la participación ciudadana y planeación del desarrollo de los departamentos.
- Generar indicadores de desarrollo institucional, que permitan clasificar las organizaciones en un momento dado y comparar su evolución a lo largo del tiempo.

1.2 Tipo de operación estadística

Es una operación estadística que se realiza por muestreo probabilístico estratificado.

1.3 Universo

Servidores públicos pertenecientes al nivel central de las gobernaciones del territorio nacional y del Distrito Capital (Bogotá D.C.).

1.4 Población objetivo

Son los servidores públicos territoriales, con una antigüedad superior a seis meses, que laboran en la administración central del Departamento/ Distrito Capital.
En total 32 gobernaciones y Distrito Capital

1.5. Variables de interés

- De estudio: percepción del funcionario frente al ambiente y desempeño institucional en términos de credibilidad en las reglas, en las políticas y frente a la suficiencia de recursos,

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

gestión por resultados, rendición de cuentas, bienestar laboral y prevención de prácticas irregulares.

- De clasificación por: entidades, nivel de cargo de los funcionarios, sexo y tiempo de servicio.

1.6. Parámetros a estimar

En los informes publicados, se proporcionan dos tipos de indicadores:

El primer tipo son dos estadísticos que evalúan las respuestas a las preguntas individuales. Uno de los estadísticos consiste en la distribución de frecuencias relativas porcentuales de las respuestas y el otro es un promedio aritmético basado en puntajes de favorabilidad asignados a cada respuesta posible.

El segundo tipo de indicador intenta evaluar las respuestas de un conjunto de preguntas que tienen en común un tema específico (por ejemplo el ambiente institucional). Este tipo de indicadores sintéticos son prácticamente los mismos que los de preguntas individuales. En el caso de la distribución de la frecuencia relativa porcentual, la diferencia radica en la clasificación que se le da a las posibles respuestas. Pues cada respuesta, dependiendo del contexto de la pregunta se le clasifica en cuatro niveles de favorabilidad: Totalmente favorable, favorable, desfavorable, totalmente desfavorable.

En el caso del promedio se construye el promedio de promedios de preguntas individuales que es igual al promedio aritmético de los puntajes de favorabilidad para todas las respuestas de todas las preguntas del tema en específico.

Los parámetros estimados vienen acompañados del coeficiente de variación estimado (cve%) y el intervalo de confianza de la estimación (ic (+-)).

1.7. Cobertura

Por entidad territorial y por total gobernaciones y Distrito Capital.

1.8. Niveles de desagregación

El diseño muestral se realizó para dar estimaciones con niveles de confianza y precisión útiles. Coeficientes de variación menores del 10% para totales y razones estimados a nivel de Entidad, cualquier otra desagregación está sujeta a que las estimaciones no sean tan precisas y su uso dependerá del coeficiente de variación estimado que presenten.

1.9. Unidades estadísticas

La unidad de observación son los servidores públicos con una antigüedad superior a seis meses en la entidad, y que laboran en el nivel central de las gobernaciones y el Distrito Capital.



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 5
FECHA: 29-08-2013

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

La unidad de análisis es la entidad pública territorial y el total gobernaciones y Distrito Capital.

Finalmente, la unidad de muestreo corresponde a los servidores públicos que pertenecen a las entidades objeto de estudio.

1.10 Periodo de referencia

La recolección de la información contenida en el cuestionario se refiere al periodo transcurrido durante los últimos doce meses.

1.11. Periodo de recolección

La información se recolecta anualmente durante un mes.



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 6
FECHA: 29-08-2013

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

2. Diseño muestral

2.1 Marco muestral

Para la encuesta sobre ambiente y desempeño institucional departamental, el marco muestral corresponde al listado de servidores públicos de cada una de las entidades objeto de estudio. Esta información es solicitada, todos los años, por las territoriales del DANE, a las oficinas de recursos humanos de las entidades participantes y al final la información se consolida en el DANE CENTRAL.

El marco contiene variables que identifican a los servidores: nombres, apellidos y cedula de ciudadanía. Otras variables que facilitan su ubicación como: la entidad, el área o dependencia, el municipio donde trabaja, la sede y el correo electrónico. Y otras variables que los caracterizan como: el cargo, el sexo y la fecha de ingreso del funcionario.

Para identificar problemas de sobre cobertura o sub cobertura del marco, cada año se compara el número de servidores de las entidades con respecto al del año anterior, cuando existen diferencias significativas se procede a confirmar y depurar la información con ayuda de la fuente.

Es poco frecuente pero puede suceder que se repitan servidores en el marco, los funcionarios que se identifican con este problema pueden estar en dos entidades diferentes o en la misma entidad. Cuando el funcionario está en dos entidades es usualmente es porque el funcionario cambió recientemente de trabajo, o se encuentra en comisión; dado que ningún funcionario debe trabajar al mismo tiempo en más de una entidad.

Con respecto a las razones para que un funcionario se repita en la misma entidad generalmente son, porque el funcionario tiene un ascenso, porque está en alguna especie de comisión o por algún error de digitación. En todo caso los servidores repetidos del marco se identifican y depuran de la forma más apropiada según sea el caso.

2.2 Tipo de muestreo

El diseño muestral propuesto para esta encuesta es estratificado y el método de selección en cada estrato es Muestreo Aleatorio Simple (MAS). El criterio de estratificación dentro de cada entidad está dado por la jerarquía del cargo, formándose tres niveles:

- Nivel 1: directivos y asesores
- Nivel 2: profesionales y técnicos
- Nivel 3: asistenciales

La estratificación busca garantizar que los resultados tomen en cuenta el punto de vista de cada nivel jerárquico. El número total de estratos depende del número de entidades y el número de niveles jerárquicos dentro de cada entidad.

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

2.3 Definición del tamaño de la muestra

El ideal de la encuesta es tomar el punto de vista de todos los funcionarios de las entidades de estudio, sin embargo existen entidades cuya población excede posibilidades logísticas y económicas. Por tal motivo, para algunas entidades se hace una muestra de funcionarios y en otras se hace censo.

Cuando se hace muestra, el número de funcionarios seleccionados, se basa en ejercicios teóricos-prácticos y experiencias en años anteriores que aseguran indicadores con coeficiente de variación menor del 5% dentro de cada entidad.

Cuando se hace censo es porque el número de funcionarios en la entidad es menor o igual a 110. Adicionalmente, el censo en las entidades que tienen menor número de servidores públicos, pretende garantizar la confidencialidad y reserva estadística de las fuentes. La relación entre tamaño poblacional de la entidad y la muestra se presenta en la siguiente tabla.

Tabla 1. Relación entre tamaño poblacional y muestra seleccionada

Número de funcionarios entre			Tamaño de muestra aproximado
Entre	111	Y 150	110
Entre	151	Y 200	120
Entre	201	Y 250	130
Entre	251	Y 350	140
Entre	351	Y 450	160
Entre	451	Y 800	230
Entre	801	Y 2000	300
Más de	2000		350

Fuente: DANE.

Con base al tamaño de muestra por entidad definido en la tabla anterior se asigna un tamaño de muestra para cada nivel jerárquico. La asignación del tamaño de muestra dentro de cada nivel es proporcional al tamaño del nivel con respeto al total de la entidad. Si cualquier nivel jerárquico tiene menos de 8 servidores, entonces se hace censo de este nivel y no muestra.

En el Anexo A se encuentran los tamaños de muestra a nivel de entidad para la EDID 2013.

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

2.4 Procedimiento de estimación

Estimadores y factor de expansión:

Los principales parámetros a estimar (indicadores simples y sintéticos) son distribuciones de frecuencias absolutas, relativas y promedios. Todos estos parámetros tienen forma de totales y de razones entre totales, particularmente las frecuencias absolutas son totales, y las frecuencias relativas y promedios son razones entre totales. El estimador utilizado para estimar totales es el de Horvitz-Thompson del cual hace parte el factor de expansión. El factor de expansión es un valor numérico asociado a cada elemento de la muestra. Su objetivo es hacer que la información del elemento seleccionado represente en el estimador del total de Horvitz-Thompson, parte del conjunto de elementos que no fueron seleccionados en la muestra. La construcción del factor de expansión no es algo fortuito sino que es el resultado del diseño de muestreo elegido; se define por el inverso de la probabilidad de inclusión del elemento seleccionado en la muestra.

En este texto, se presenta la probabilidad de inclusión del elemento k como π_k , y el valor de la variable y para el elemento k como y_k . El estimador Horvitz-Thompson del total poblacional de la variable y en una muestra s se define entonces como:

$$\hat{t}_{y\pi} = \sum_s \frac{1}{\pi_k} y_k$$

Por ejemplo, cuando se tiene un diseño con muestreo aleatorio simple, la probabilidad de inclusión del elemento k toma la forma de $\pi_k = \frac{n}{N}$. Donde n es el tamaño de la muestra y N es el Tamaño de la población.

Existen dos tipos de particiones de la población que se toman en cuenta en el cálculo de los estimadores. Una partición crea grupos llamados estratos, y la otra crea grupos llamados dominios. La diferencia conceptual entre ambas particiones de la población radica en la función que desempeñan. La partición en estratos asigna un diseño muestral independiente a cada estrato induciendo muestras y probabilidades de inclusión particulares por estrato s_h y π_{hk} respectivamente. Por otro lado, la partición en dominios posibilita generar resultados por cualquier desagregación de interés temático.

En un diseño estratificado en H estratos el estimador del total poblacional, de la variable y , toma la forma de:

$$\hat{t}_{y\pi_H} = \sum_{h=1}^H \hat{t}_{y\pi_h} = \sum_{h=1}^H \sum_{s_h} \frac{1}{\pi_{hk}} y_k$$

Donde $\hat{t}_{y\pi_h}$ es el estimador del total de la variable y para la población del estrato h . En el caso de un diseño con muestreo aleatorio simple estratificado, la probabilidad de

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

inclusión del elemento k en el estrato h toma la forma de $\pi_{hk} = \frac{n_h}{N_h}$ Donde n_h es el tamaño de la muestra en el estrato h y N_h es el Tamaño de la población del estrato h . Para un dominio d de la población en una muestra s , el total de la variable y toma la forma de:

$$\hat{t}_{y_d\pi} = \sum_s \frac{1}{\pi_k} y_k z_{dk}$$

Con $z_{dk} = 1$ si el elemento k pertenece al dominio d y $z_{dk} = 0$ si no lo hace.

Si se desea estimar un dominio d bajo un diseño estratificado en H estratos el estimador del total poblacional, de la variable y toma la forma de:

$$\hat{t}_{y_d\pi_H} = \sum_{h=1}^H \hat{t}_{y_d\pi_h}$$

Donde $\hat{t}_{y_d\pi_h}$ es el estimador del total de la variable y en el dominio d para la población del estrato h .

El otro parámetro típico estimado para la encuesta es el parámetro de la razón entre dos totales. Sean dos variables de interés y y w . El estimador de la razón entre los totales de las variables y y w con una muestra s , se define como:

$$\hat{R} = \frac{\hat{t}_{y\pi}}{\hat{t}_{w\pi}}$$

Nótese que tanto el numerador como el denominador son estimadores de totales de Horvitz-Thompson. Para el diseño estratificado en H estratos el estimador de la razón entre totales, de las variables y y w , toma la forma de:

$$\hat{R}_H = \frac{\hat{t}_{y\pi_H}}{\hat{t}_{w\pi_H}} = \frac{\sum_{h=1}^H \hat{t}_{y\pi_h}}{\sum_{h=1}^H \hat{t}_{w\pi_h}}$$

Para un dominio d de la población en una muestra s el estimador de la razón entre totales, de las variables y y w , toma la forma de:

$$\hat{R}_d = \frac{\hat{t}_{y_d\pi}}{\hat{t}_{w_d\pi}}$$

Si se desea estimar dominio d , en un diseño estratificado en H estratos el estimador de la razón entre totales, de las variables y y w , toma la forma de:

$$\hat{R}_{dH} = \frac{\hat{t}_{y_d\pi_H}}{\hat{t}_{w_d\pi_H}} = \frac{\sum_{h=1}^H \hat{t}_{y_d\pi_h}}{\sum_{h=1}^H \hat{t}_{w_d\pi_h}}$$

La forma más general de expresar estimadores de total y de razón son los $\hat{t}_{y_d\pi_H}$ y \hat{R}_{dH}

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

respectivamente, porque las formulas de estimadores $\hat{t}_{y\pi}$, $\hat{t}_{y\pi_H}$, \hat{R} y \hat{R}_H son un caso particular de $\hat{t}_{y\pi_H}$ y \hat{R}_{dH} , solo es necesario definir la variable dominio (z_{dk}) de manera apropiada. Por tal motivo en la descripción que se hace mas adelante de los indicadores, se va a tomar la forma general de estimación, es decir, totales con $\hat{t}_{y\pi_H}$ y razones por \hat{R}_{dH} . Cada estimador calculado con la probabilidad inclusión del elemento k por $\pi_{hk} = \frac{n_h}{N_h}$ asociada a un diseño con muestreo aleatorio simple estratificado.

Corrección por no respuesta:

Existen dos tipos de no respuesta de los elementos de la muestra. El primero es la no respuesta total y se da cuando ninguna de las variables posee información. El segundo es la no respuesta parcial y se da cuando por lo menos una de las variables tiene información.

Con respecto a la no respuesta total, se evidencia que no hay factores particulares dentro de los estratos de muestreo que la ocasionen, así que se asume que la no respuesta es aleatoria dentro del estrato. Por esta situación la corrección se hace a través de un factor de ajuste que multiplica al factor de expansión. El factor se construye por estrato y toma en cuenta lo siguiente:

- *Elementos fuera del universo:* son todos aquellos elementos seleccionados en la muestra que no pertenecen al universo de estudio; entre éstos se encuentran servidores públicos que ya no pertenecen a la entidad por causa de defunción, jubilación o retiro.
- *Elementos sin información:* corresponde a los elementos que pertenecen al universo de estudio, pero no se obtiene información de ellos. Dentro de este grupo se encuentran los servidores que no se pudieron localizar durante el operativo de campo y los que rechazaron realizar la encuesta.

Con lo cual se define el factor de ajuste para la no respuesta total como:

$$F_{ajust(h)} = \frac{n_h - n_{fuera\ del\ universo(h)}}{n_h - n_{no\ respuesta(h)} - n_{fuera\ del\ universo(h)}}$$

Donde

$n_{fuera\ del\ universo(h)}$: Total servidores públicos seleccionados en el estrato h que no pertenecen al universo de estudio.

$n_{no\ respuesta(h)}$: Total servidores públicos seleccionados que, aunque pertenecen al universo de estudio en el estrato h no responden el cuestionario.

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

Con respecto a la no respuesta parcial, el tratamiento es imputar los datos a través del método de paquete caliente. Cabe resaltar que los procedimientos para la recolección de la información han minimizado las tasas de no respuesta dentro de los estratos haciendo que sean pocos los datos que se imputen. En la imputación se buscan donantes que tengan las mismas características, particularmente el donante tomado en esta encuesta tiene que pertenecer a la misma entidad y ocupar el mismo nivel de cargo del receptor.

Cálculo de precisión de los resultados

La precisión de la estimación se construye utilizando la varianza estimada del estimador del parámetro. Para el caso con muestreo aleatorio simple estratificado, el estimador de la varianza del estimador del total de la forma general $\hat{t}_{y_d\pi_H}$ es:

$$\widehat{VAR}(\hat{t}_{y_d\pi_H}) = \sum_{h=1}^H \frac{N_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right) S_{hy_d}^2$$

$$S_{hy_d}^2 = \frac{\sum_{s_h} (y_{dk} - \bar{y}_d)^2}{n_h - 1}$$

Donde $S_{hy_d}^2$: Varianza muestral de la variable y con dominio d , simbolizada como y_d , de la muestra s_h sacada del estrato h . Por otro lado, para construir el estimador de la varianza del estimador de la razón de la forma general \hat{R}_{dH} (ver arriba), entre los totales de las variables y y w con dominio d , se debe primero calcular una nueva variable u_d para cada elemento k :

$$u_{dk} = \frac{1}{\hat{t}_{w_d\pi_H} (y_{dk} - \hat{R}_{dH} w_{dk})}$$

Donde $y_{dk} = y_k z_{dk}$ y $w_{dk} = w_k z_{dk}$ con $z_{dk} = 1$ si el elemento k pertenece al dominio d y $z_{dk} = 0$ si no lo hace. El estimador de la varianza de la razón es equivalente al estimador de la varianza del estimador del total de la nueva variable u_d :

$$\widehat{VAR}(\hat{t}_{u_d\pi_H}) = \widehat{VAR}(\hat{R}_{dH}) = \widehat{VAR}\left(\frac{\hat{t}_{y_d\pi_H}}{\hat{t}_{w_d\pi_H}}\right) = \sum_{h=1}^H \frac{N_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right) S_{hu_d}^2$$

Donde $S_{hu_d}^2$: Varianza muestral de la variable u en el estrato h con dominio d

Con la varianza estimada del total o de la razón, se puede construir el coeficiente de variación estimado del parámetro como medida de precisión, la idea es que entre menor sea el coeficiente de variación estimado cve, menor incertidumbre habrá sobre la estimación, su fórmula es:

$$cve = 100 * \frac{\sqrt{\widehat{VAR}(\hat{\theta})}}{\hat{\theta}}$$

Donde $\hat{\theta} = \hat{t}_{y_d\pi_H}, \hat{R}_{dH}$ corresponde al valor estimado del parámetro de interés. También se puede definir un intervalo de confianza del 95%:

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

$$IC(\hat{\theta}) = \hat{\theta} \pm 1.96 \sqrt{VAR(\hat{\theta})}$$

Interpretación de la precisión

Uno de los principales criterios para determinar la calidad de la estimación de un parámetro es la variabilidad que tienen los posibles resultados de dicha estimación. Esta variabilidad se conoce como varianza del estimador, la cual depende de muchos factores, como el diseño muestral, el tamaño de la muestra, el parámetro que se desea estimar, los niveles de desagregación, entre otros.

La varianza se calcula básicamente como la suma del cuadrado de las distancias entre los diferentes valores de una variable y su valor promedio, en este caso, los diferentes valores corresponden a las posibles estimaciones, las cuales, a su vez, provienen de las posibles muestras. Posteriormente, se toma el cuadrado de la distancia para evitar que éstas se anulen entre sí y se disfrace la verdadera dispersión de los datos; por esta razón, la varianza proporciona la variabilidad en unidades al cuadrado, es decir, personas al cuadrado, ganado al cuadrado o hectáreas cuadradas, lo que no permite una comprensión fácil de esta magnitud.

La raíz cuadrada de esta varianza es la que se denomina desviación estándar de la distribución o error estándar. Esta medida de dispersión tiene la ventaja que la unidad de medida de dispersión corresponde a la unidad de la variable de interés, se establece en términos de personas, hectáreas o pesos, aunque queda la dificultad de saber si una desviación es grande o pequeña; así, por ejemplo, una variabilidad de un millón de pesos puede ser muy grande si se habla del promedio de ingresos de los empleados, pero es absolutamente pequeño si se determina sobre el total del volumen de ventas en la industria del país.

El coeficiente de variación estimado, más conocido como error de muestreo, corresponde al valor que indica el grado de precisión con el cual se está reportando un resultado de las estimaciones de los parámetros definidos con anterioridad. Es decir, se trata de la magnitud de la incertidumbre de una estimación. Se define como la variación porcentual del error estándar a la estimación central, es decir, se trata del cociente entre el error estándar del estimador y el estimador multiplicado por 100, así:

$$CV = 100 * \frac{\sqrt{VAR(\hat{\theta})}}{E(\hat{\theta})}$$

Y el coeficiente de variación estimado está dado por:

$$cve = 100 * \frac{\sqrt{VAR(\hat{\theta})}}{\hat{\theta}}$$

Aunque la varianza, el error estándar y el coeficiente de variación miden la magnitud de la variabilidad de la distribución muestral del estimador, es decir, lo que comúnmente se



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 13
FECHA: 29-08-2013

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

denomina error de muestreo, el coeficiente de variación tiene la ventaja de proporcionar esta medida en términos porcentuales, por ello se constituye en una medida común para estimaciones.

Se suele considerar que el resultado de una estimación es bueno si su coeficiente de variación es menor de 5%; aceptablemente práctico, entre 5% y 10%; de baja precisión si está en el rango mayor de 10% y menor de 15%; y no útil si es mayor a 15%.

Para entender mejor el significado y los diferentes valores que toman los coeficientes de variación en los cuadros presentados, se debe tener en cuenta que el diseño de la muestra se realizó para obtener estimaciones con alta precisión a nivel nacional por entidad. Las estimaciones para otros niveles de desagregación (como sector) están sujetas a que su precisión no necesariamente sea buena y por tanto el dato no sea confiable.

Es por esta razón que en algunos cuadros aparece, por ejemplo, el total de una variable a nivel nacional y por alguna categoría de análisis con coeficientes de variación pequeños, mientras que para otras categorías de la misma variable los cve son muy altos, en ocasiones del 30 % e incluso mayores del 100 %. En estos casos, el DANE publica la cifra aunque no sea confiable, básicamente para que en los cuadros de salida la información de los totales se observe consistente y porque en muchos casos, el usuario por operaciones aritméticas simples puede deducir el valor correspondiente a esa estimación. Sin embargo, es muy importante que los usuarios de la información sean conscientes del bajo nivel de precisión que tienen estas estimaciones.

Esta situación puede darse por varias causas. Por ejemplo, cuando el fenómeno estudiado ocurre con gran frecuencia en algunas de las categorías de la variable de clasificación, por lo cual la estimación para dichas categorías es de alta calidad; pero puede ocurrir que para otras categorías en las que el fenómeno no es frecuente, la estimación no es buena, pues el tamaño de muestra no es suficiente. También puede obedecer al hecho de que en algunas de estas categorías el fenómeno es muy variable mientras en otras es más uniforme, lo que genera menor varianza en las estimaciones.

Todos los resultados que se producen se presentan en cuadros de salida, donde cada estimación tiene su respectivo cve o error muestral. En general, a medida que se incrementa la desagregación de las estimaciones, según las variables de clasificación, el error muestral se incrementa, básicamente porque el tamaño de muestra que incide sobre estos grupos es menor

PROCESO: Diseño

SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID

ELABORÓ: Diseños Muestrales

REVISÓ: Coordinación Temática Social

APROBÓ: Director Técnico DIMPE

3. Diseño de indicadores

Los indicadores de la encuesta resumen dos tipos de información, la información que se encuentra por variable y la información que se tiene por conjunto de variables. A los primeros indicadores se les llama indicadores simples y a los segundos, indicadores sintéticos.

3.1 INDICADORES SIMPLES

Las variables de la encuesta son de tipo categórico y cada variable de estudio cuenta con cuatro posibles categorías, por tal motivo se recurre a la distribución de frecuencias relativas porcentuales para la presentación de resultados por variable.

Otro indicador simple que se tiene para la encuesta es el promedio aritmético, el cual se construye a través de la transformación de las variables categóricas en variables discretas. La transformación se hace a partir de un puntaje que se asigna temáticamente a cada categoría. La idea de este puntaje es caracterizar cada categoría en cuatro posibles valores numéricos, 1 cuando la variable toma una categoría totalmente desfavorable, 2 cuando es desfavorable, 4 cuando es favorable y 5 si es totalmente favorable. Todos estos valores miden la favorabilidad de las respuestas de los funcionarios hacia las entidades donde laboran.

3.2 INDICADORES SINTÉTICOS

Los indicadores sintéticos intentan resumir la información contenida en un conjunto de variables que tienen en común un concepto temático. Para construir los indicadores es necesario homogeneizar la escala de las categorías de cada variable, es decir todas las variables deben tener las mismas categorías.

Las variables que forman parte del indicador grupal tienen cuatro categorías, que se pueden asociar a cuatro posibles niveles de favorabilidad del servidor público hacia su entidad. Nivel 1: Totalmente desfavorable, Nivel 2: desfavorable, Nivel 3: favorable, Nivel 4: totalmente favorable. Con estas cuatro posibles categorías se construye la distribución de frecuencias relativas porcentuales uniendo todas las variables en una sola.

Así mismo, estas categorías se transforman en valores numéricos, que cuantifican el nivel de favorabilidad; 1 para totalmente desfavorable, 2 para desfavorable, 4 para favorable y 5 para totalmente favorable. Con esta transformación se puede construir el promedio de promedios de las variables como otro indicador sintético.

3.3 CÁLCULO DE INDICADORES SIMPLES

Los indicadores simples están definidos para variables individuales. Se estima la frecuencia absoluta y relativa de cada categoría y el promedio con la asociación numérica

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

presentada en la sección anterior. A continuación se presentarán de forma general los indicadores simples y su respectiva forma de estimación.

Para el conjunto poblacional U de N elementos sea y una variable categórica con categorías $c_1, c_2, \dots, c_i, \dots, c_I$. El indicador de frecuencia absoluta de la categoría c_i mide el número de elementos que están clasificados en esta categoría. El indicador se calcula como:

$$t_{c_i} = \sum_U z_{c_i,k}$$

Con $z_{c_i,k} = 1$ si el elemento k se clasifica en la categoría c_i y $z_{c_i,k} = 0$ si se clasifica en otra categoría. Se observa que el indicador frecuencia absoluta es el total de la variable z_{c_i} por lo tanto este total se estima de la forma general $\hat{t}_{z_{c_i}}$.

Por otro lado el indicador de frecuencia relativa de la categoría c_i mide el porcentaje de elementos que están clasificados en esta categoría. El indicador se calcula para el conjunto poblacional U de N elementos como:

$$R_{c_i} = \frac{t_{c_i}}{N} = \frac{\sum_U z_{c_i,k}}{\sum_U 1}$$

Se observa que el indicador de frecuencia relativa es la razón de dos totales. El numerador es el indicador de la frecuencia absoluta de la categoría c_i y denominador es el total de sumar una variable de valor constante de unos. Entonces esta razón se estima de la forma general \hat{R}_{dH} .

El indicador de promedios para la variable categórica y , anteriormente definida, utiliza la transformación de las I categorías en valores numéricos. La transformación genera una nueva variable y' sobre la cual se calcula el promedio poblacional. La fórmula para este indicador es:

$$P = \frac{t'_y}{N} = \frac{\sum_U y'_k}{\sum_U 1}$$

Al igual que el indicador de frecuencias relativas, el indicador de promedio es una razón entre dos totales, por lo cual se estima a través de la forma general \hat{R}_{dH} .

3.4 CÁLCULO DE INDICADORES SINTÉTICOS

Los indicadores sintéticos están definidos para conjuntos de variables. Se estiman las frecuencias relativas de los niveles de favorabilidad y promedios a través de la relación numérica de los niveles. A continuación se presentarán de forma general los indicadores sintéticos y su respectiva forma de estimación.

Para el conjunto poblacional U de N elementos con J variables categóricas, todas con las mismas I categorías $c_1, c_2, \dots, c_i, \dots, c_I$. Se define el indicador de frecuencia absoluta de

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

la categoría c_i del conjunto de N elementos y J variables como:

$$T_{c_i} = \sum_{j=1}^J \sum_U z_{c_i k j}$$

Con $z_{c_i k j} = 1$ si el elemento k se clasifica en la categoría c_i de la variable j y $z_{c_i k j} = 0$ si se clasifica en otra categoría para cualquier variable j . El indicador también se puede

escribir sobre la suma de una variable nueva $Z_{c_i} = \sum_{j=1}^J z_{c_i k j}$ de la siguiente forma:

$$T_{c_i} = \sum_{j=1}^J \sum_U z_{c_i k j} = \sum_U \sum_{j=1}^J z_{c_i k j} = \sum_U Z_{c_i k}$$

Por lo tanto el indicador sintético de frecuencia absoluta es el total de la variable Z_{c_i} y se estima de la forma general $\hat{t}_{Z_{dH}}$.

Por otro lado el indicador de frecuencia relativa de la categoría c_i mide el porcentaje de valores que están clasificados en esta categoría para el conjunto de las J variables. El indicador se calcula como:

$$R_{c_i} = \frac{T_{c_i}}{NJ} = \frac{\sum_U z_{c_i k}}{\sum_U J}$$

El cual es una razón de totales de dos variables; en el denominador la frecuencia absoluta y en el denominador el total de una variable de valor constante J . Entonces esta razón se estima de la forma general \hat{R}_{dH} .

El indicador de promedios para el conjunto de variables categóricas y_j utiliza la transformación de las J categorías en valores numéricos para las J variables. La transformación genera un conjunto de nuevas variables y'_j sobre las cuales se calcula el promedio poblacional y luego se promedian los promedios. La fórmula para este indicador es:

$$P = \frac{1}{J} \sum_{j=1}^J \frac{t_{y'_j}}{N}$$

El promedio de promedios se puede reescribir como:

$$P = \frac{1}{J} \sum_{j=1}^J \frac{t_{y'_j}}{N} = \frac{\sum_{j=1}^J t_{y'_j}}{NJ} = \frac{\sum_{j=1}^J \sum_U y'_{kj}}{NJ} = \frac{\sum_U \sum_{j=1}^J y'_{kj}}{NJ} = \frac{\sum_U Y'_k}{NJ} = \frac{\sum_U Y'_k}{\sum_U J}$$

$$Y' = \sum_{j=1}^J y'_j$$

Lo cual es una razón entre los totales de la variable nueva Y' y la variable de valor constante J . La forma de estimar este indicador se hace a través de la forma general \hat{R}_{dH} .

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID		
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE	

3.5 TIPOS DE VARIABLES

Las variables de estudio son categóricas y tienen por lo general 4 categorías. Conceptualmente no existen categorías intermedias, para obligar al encuestado a tomar una posición positiva o negativa acerca de las preguntas que se le hacen. Se tienen las siguientes posibles opciones de respuesta:

Totalmente de acuerdo, De acuerdo, En desacuerdo, Totalmente en desacuerdo.
Mucho, Algo, Poco, Nada.
Muy efectiva, Algo efectiva, Poco efectiva, Nada efectiva.
Frecuentemente, Algunas veces, Rara vez, Nunca.

A estas categorías se les puede asociar un nivel de favorabilidad que refleja la percepción favorable o desfavorable del funcionario con respecto a algunas situaciones que se presentan en la entidad a la que pertenece. Esta asociación se hace de forma numérica y también categórica como se muestra en la siguiente tabla:

Tabla 2. Asociación numérica por opción de respuesta

Asociación categórica	Totalmente favorable	Favorable	Desfavorable	Totalmente desfavorable
Asociación Numérica	5	4	2	1

Fuente: DANE.

La asociación numérica da la capacidad de cuantificar con un puntaje de uno a cinco las mediciones categóricas de cada variable, con lo cual se puede construir el promedio de la favorabilidad por variable o por conjunto de variables. Con la asociación categórica se puede observar la distribución de la favorabilidad para un conjunto de variables o una por una.

3.6 CONFORMACIÓN DE LOS INDICADORES

Indicador ambiente institucional

El indicador de ambiente institucional se define como la percepción de los funcionarios con respecto a las reglas, políticas y recursos disponibles en la entidad. A través de este indicador se caracteriza la calidad de los recursos institucionales, físicos, humanos y financieros puestos a disposición de la entidad.

Las dimensiones que conforman este concepto son:

- Credibilidad en las reglas
- Credibilidad en las políticas
- Suficiencia de recursos y previsibilidad



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 18
FECHA: 29-08-2013

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

Indicador desempeño institucional

Definido como la percepción de los funcionarios con respecto a la capacidad de la entidad para emprender procesos relacionados con la gestión por resultados, la rendición de cuentas, la promoción del bienestar laboral y la prevención de prácticas irregulares.

Las componentes que conforman este concepto son:

- Gestión por resultados
- Rendición de cuentas
- Bienestar laboral
- Prevención de prácticas irregulares



Diseño Estadístico
Encuesta sobre Ambiente y Desempeño Institucional
Departamental
-EDID-

CÓDIGO: DSO-EDID-DES-01
VERSIÓN: 01
PÁGINA: 19
FECHA: 29-08-2013

PROCESO: Diseño

SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID

ELABORÓ: Diseños Muestrales

REVISÓ: Coordinación Temática Social

APROBÓ: Director Técnico DIMPE

BIBLIOGRAFÍA

Bautista S., Leonardo, Diseños de Muestreo Estadístico, Universidad Nacional de Colombia , Departamento de Matemáticas y Estadística, (1998)

Särndal, Carl Erik, Swenson, B. ,Wretman, J. Model Assisted Survey Sampling, Springer-Verlang, New York (1992)

PROCESO: Diseño	SUBPROCESO: Encuesta sobre Ambiente y Desempeño Institucional Departamental - EDID	
ELABORÓ: Diseños Muestrales	REVISÓ: Coordinación Temática Social	APROBÓ: Director Técnico DIMPE

Anexo A. EDID 2013. Tamaño de muestra por entidad

Entidad	Población ¹	Muestra	Nrecogido	DIF
GOBERNACIÓN DE ANTIOQUIA	1750	307		
GOBERNACIÓN DEL AMAZONAS	100	100		
GOBERNACIÓN DEL ATLÁNTICO	386	191		
GOBERNACIÓN DE ARAUCA	123	110		
BOGOTÁ, D.C. ²	3649	1846		
GOBERNACIÓN DE BOLÍVAR	274	159		
GOBERNACIÓN DE BOYACÁ	466	230		
GOBERNACIÓN DE CALDAS	314	177		
GOBERNACIÓN DE CAQUETÁ	189	126		
GOBERNACIÓN DE CASANARE	246	149		
GOBERNACIÓN DEL CAUCA	319	179		
GOBERNACIÓN DEL CESAR	236	147		
GOBERNACIÓN DE CÓRDOBA	236	145		
GOBERNACIÓN DE CUNDINAMARCA	1141	300		
GOBERNACIÓN DEL CHOCÓ	44	44		
GOBERNACIÓN DEL HUILA	229	143		
GOBERNACIÓN DE GUAINÍA	40	40		
GOBERNACIÓN DE LA GUAJIRA	155	119		
GOBERNACIÓN DEL GUAVIARE	101	101		
GOBERNACIÓN DEL MAGDALENA	188	125		
GOBERNACIÓN DEL META	427	199		
GOBERNACIÓN DE NARIÑO	235	144		
GOBERNACIÓN DE NORTE DE SANTANDER	283	162		
GOBERNACIÓN DEL PUTUMAYO	111	110		
GOBERNACIÓN DEL QUINDÍO	280	160		
GOBERNACIÓN DE RISARALDA	289	164		
GOBERNACIÓN DE SAN ANDRÉS	226	142		
GOBERNACIÓN DE SANTANDER	387	197		
GOBERNACIÓN DE SUCRE	299	162		
GOBERNACIÓN DE TOLIMA	282	162		
GOBERNACIÓN DEL VALLE DEL CAUCA	809	300		
GOBERNACIÓN DEL VAUPÉS	101	101		

Fuente: DANE

¹ Para poblaciones inferiores a 110 servidores públicos se encuestan en su totalidad.

² En la muestra para Bogota se incluyen la información correspondiente a la Alcaldía Mayor, Secretarías Distritales de Ambiente, Secretaría de Educación Distrital, Secretaría Distrital de Cultura, Recreación y Deporte; Secretaría Distrital de Gobierno, Secretaría Distrital de Hacienda, Secretaría Distrital de Integración Social; Secretaría Distrital de Movilidad; Secretaría Distrital de Planeación, Secretaría Distrital de Salud y Secretaría Distrital del Hábitat y Secretaria Distrital de Desarrollo Económico.