

DISEÑO DE LA MUESTRA¹

1. POBLACION DEL ESTUDIO

La población de estudio está constituida por todos los proyectos de investigación financiados por COLCIENCIAS que finalizaron entre el año de 1999 y 2005.

Como población de referencia, se consideran los listados de proyectos que fueron suministrados por COLCIENCIAS. En la tabla siguiente se describe el número de proyectos relacionados en las dos bases de datos recibidas. La Base 1 contiene los proyectos cuyo tipo de financiación es Recuperación Contingente RC, Cofinanciación COF, línea de crédito BANCOLDEX y Recuperación Contingente de Regionalización. En la Base 2 se encuentran los proyectos que recibieron Incentivos Tributarios. De estas dos bases se eliminaron 392 proyectos, 127 porque tuvieron algunas observaciones de los directores de los Programas de COLCIENCIAS y 265 que recibieron incentivos tributarios en los años 1999 y 2000 (119 en el año 1999 y 146 en el 2000), debido a que no tenían la información sobre el programa de COLCIENCIAS al cual pertenece. De esta manera se logró tener un total de 2.840 que podrán constituir el marco de referencia del estudio.

Tabla 1. Número de proyectos en las bases de datos de COLCIENCIAS.

AÑO	BASE 1	BASE 2	TOTAL BASES	ELIMINADOS	TOTAL
1999	283	119	402	147	255
2000	406	146	552	146	406
2001	182	39	221		221
2002	459	121	580		580
2003	444	85	529	38	491
2004	416	99	515	36	479
2005	329	104	433	25	408
TOTAL	2519	713	3232	392	2840

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.

Una vez depuradas las bases de datos atendiendo los términos de referencia del proyecto y las observaciones formuladas por cada uno de los jefes de Programas de COLCIENCIAS, se seleccionaron los proyectos que integran la población de estudio de acuerdo con los siguientes criterios:

¹ Documento elaborado por Jorge Martínez, equipo Universidad Nacional de Colombia.

- a. Tener en la base de datos los datos requeridos para hacer la estratificación.
- b. No tener observaciones de exclusión por parte de los directores de Programas de COLCIENCIAS.
- c. Tener fecha de vencimiento anterior a diciembre 31 del año 2005 y en el caso de que en la base de datos figure fecha de prórroga, esta sea anterior al 31 de diciembre de 2005.
- d. Si en la base de datos no existe fecha de vencimiento o fecha de prórroga, el año más los meses de plazo registrados no deben superar el año 2005.

Los 1922 proyectos que cumplieron estos criterios se registran en la tabla siguiente, junto con el número de proyectos que no fueron eliminados en la primera revisión, los cuales aparecen además en la tabla anterior.

Tabla 2. Relación anual de los proyectos del marco de referencia y de la población de estudio.

AÑO	TOTAL	POBLACION
1999	255	255
2000	406	406
2001	221	214
2002	580	525
2003	491	281
2004	479	130
2005	408	111
TOTAL	2840	1922

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.

La tabla siguiente muestra la distribución de la población de proyectos objeto de estudio, según los programas de COLCIENCIAS y las 3 regionales consideradas en el proyecto.

Tabla 3. Distribución de la población según los programas de COLCIENCIAS y las diferentes regionales.

PROGRAMAS DE COLCIENCIAS	REGION 1	REGION 2	REGION 3	TOTAL
AGROPECUARIAS	29	31	9	69
BASICAS	88	49	16	153
BIOTECNOLOGIA	51	33	5	89
DESARROLLO TECNOLOGI	122	79	31	232
EDUCACION	43	51	21	115
ELECTRONICA	55	85	10	150
ENERGIA	37	18	229	284
INNOVACION	138	99	46	283
MAR	8	13	16	37
MEDIO AMBIENTE	39	26	5	70
PROGRAMAS ESTRATEGIC	2	13	3	18
REGIONALIZACION	14	6	8	28
SALUD	116	108	41	265
SOCIALES	35	85	8	128
TOTAL	777	696	448	1921

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.

Uno de los proyectos no es asimilable a ninguna de las regiones, ya que el lugar de ejecución es la ciudad de Miami. De aquí que el tamaño de la población es 1.921 en lugar de los 1.922 de la tabla anterior.

Como en los términos de referencia se establece que se debe realizar un estudio independiente en cada regional, se considera entonces, que cada una de las tres Regiones es una población diferente.

2. ESTRATIFICACION DE LAS POBLACIONES

Con el fin de lograr una muestra representativa de cada población, estas se estratifican según los Programas de COLCIENCIAS y los grupos de proyectos establecidos en los términos de referencia:

- Grupo 1 Proyectos en modalidad recuperación contingente: proyectos de Investigación científica y aplicada.
- Grupo 2 Proyectos en modalidad Cofinanciación, empresa- universidad y empresa-centro de desarrollo tecnológico: proyectos de investigación y desarrollo e innovación realizados de manera conjunta.
- Grupo 3 Proyectos en modalidad de crédito línea de crédito BANCOLDEX-COLCIENCIAS: proyectos de innovación, realizados por empresas de todos los tamaños y temas.
- Grupo 4. Proyectos en modalidad de incentivos tributarios: proyectos de innovación presentados por las empresas.

Grupo 5. Proyectos y programas en modalidad de recuperación contingente de regionalización

La distribución de la población según estos criterios de estratificación se muestra en la tabla siguiente.

Tabla 4. Estratificación de la población de la Región 1 Occidente asignada a TECNOS según los Programas de COLCIENCIAS y los cinco tipos de proyectos.

PROGRAMAS DE COLCIENCIAS	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	SIN IDENT.	TOTAL
AGROPECUARIAS	12	10	0	7	0	0	29
BASICAS	71	0	0	17	0	0	88
BIOTECNOLOGIA	30	1	0	20	0	0	51
DESARROLLO TECNOLOGICO	20	48	3	51	0	0	122
EDUCACION	40	0	0	2	0	1	43
ELECTRONICA	11	21	0	23	0	0	55
ENERGIA	9	19	0	9	0	0	37
INNOVACION	138	0	0	0	0	0	138
MAR	7	1	0	0	0	0	8
MEDIO AMBIENTE	33	2	0	4	0	0	39
PROGRAMAS ESTRATEGICOS	1	0	0	1	0	0	2
REGIONALIZACION	1	0	0	1	12	0	14
SALUD	103	1	0	12	0	0	116
SOCIALES	34	0	0	1	0	0	35
TOTAL	510	103	3	148	12	1	777

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.

3. TIPO DE MUESTRA

La muestra de proyectos para el presente estudio es de carácter probabilística y estratificada. Probabilística en el sentido de que cada proyectos tendrá una probabilidad previamente establecida de pertenecer a la muestra y estratificada porque la selección de los proyectos se realiza atendiendo la estratificación definida anteriormente.

4. PRECISION Y CONFIABILIDAD

La muestra es seleccionada de tal manera que tenga un error relativo que no supere el 7%, con un nivel de confiabilidad del 95%.

Para efectos de determinar el error relativo en cada una de las poblaciones se calculó el promedio y la variación de la duración de los proyectos, característica que fue definida por el grupo de investigadores como variable de referencia para el diseño de la muestra. En la tabla siguiente se muestran las estadísticas que caracterizan esta variable en las tres poblaciones. Dado que en los proyectos de incentivos tributarios no se contó con esta información solo se usaron los datos de los proyectos de los otros 4 grupos.

Tabla 5. Estadísticas de la variable de referencia.

REGIONAL	N	Media	Varianza	CV
1	627	22.7	79.4	39.3
2	611	21.1	67.9	39.0
3	225	23.1	82.9	39.4

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.

Como puede observarse en la tabla anterior en la regional 3 se presenta una mayor variabilidad, sin embargo las tres regionales tiene igual coeficiente de variación. Este hecho tiene especial efecto en el tamaño de la muestra.

5. CALCULO DEL TAMAÑO DE LA MUESTRA

El tamaño de la muestra en cada regional se calcula mediante la expresión:

$$n_o = \frac{z^2 \sigma^2}{\varepsilon^2}$$

$$n = \frac{n_o}{1 + \frac{n_o}{N}}$$

Siendo z , el percentil de la distribución normal que garantiza el nivel de confianza

deseado.

σ^2 , la varianza de la variable de referencia

ε , el margen de error deseado

N , el tamaño de la población.

Aplicando las fórmulas anteriores se obtuvo un tamaño de muestra de 98 proyectos para la Región 1. Se prevé además una sobre muestra del 20% para reemplazar los proyectos que por alguna razón no puedan ser contactados.

Una vez calculada el tamaño de la muestra este se distribuyó proporcionalmente en cada uno de los estratos de manera proporcional a los tamaños poblacionales.

Las tablas siguientes presentan la distribución del tamaño de la muestra para cada uno de los estratos.

Tabla 6. Distribución de la muestra en la Región 1 Occidente -TECNOS.

PROGRAMAS	MUESTRA REGION 1						TOTAL
	GRUP01	GRUP02	GRUP03	GRUP04	GRUP05	GRUP06	
AGROPECUARIAS	2	2	0	1	0	0	5
BASICAS	12	0	0	2	0	0	14
BIOTECNOLOGIA	5	0	0	2	0	0	7
DESARROLLO TECNOLOGI	3	8	0	8	0	0	19
EDUCACION	7	0	0	0	0	0	7
ELECTRONICA	1	4	0	3	0	0	8
ENERGIA	1	3	0	1	0	0	5
INNOVACION	5	0	0	0	0	0	5
MAR	1	0	0	0	0	0	1
MEDIO AMBIENTE	5	0	0	0	0	0	5
PROGRAMAS ESTRATEGIC	0	0	0	0	0	0	0
REGIONALIZACION	0	0	0	0	2	0	2
SALUD	15	0	0	1	0	0	16
SOCIALES	4	0	0	0	0	0	4
TOTAL	61	17	0	18	2	0	98

Fuente. Cálculos realizados por el autor con base en los datos suministrados por COLCIENCIAS.