

# PROCESO ALEATORIO PARA INGRESO DE MUESTRA ELECTRONICA

**- ACTUALIZACIÓN 2012 -**

**ACINPRO**  
***ENERO 23 AL 25 DE 2012***

## 1. DESCRIPCION GENERAL Y ANTECEDENTES

En el año 2006, en ACINPRO se desarrolló una metodología para determinar un tamaño de muestra sobre una población de planillas físicas de reporte del uso de la música, proveniente de los radiodifusores.

En el transcurso del año 2010, se decidió aplicar esta metodología a las planillas electrónicas de reporte del uso de la música proveniente de los radiodifusores, ya que las planillas físicas entraron prácticamente en desuso.

En el mes de septiembre del año 2010, se comenzó el proceso de muestreo aleatorio para planillas electrónicas, para lo cual se contó con la valiosa colaboración del Ingeniero John Jairo Roldán. Sin embargo, posteriormente se evidenció la necesidad de realizar ajustes a la metodología, debido a que las fórmulas iniciales fueron calculadas con un margen de población que con el paso del tiempo, cambió sustancialmente.

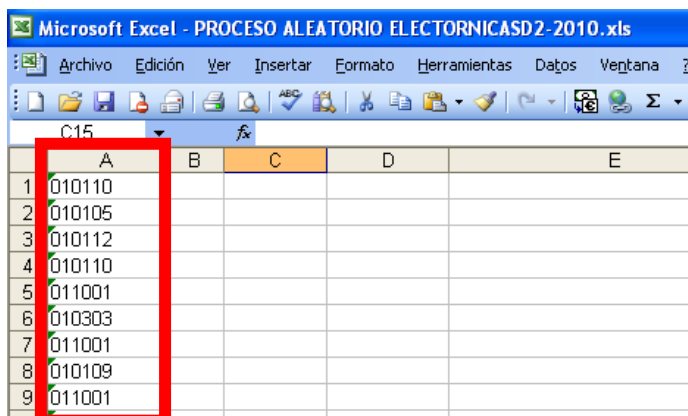
Nuevamente, con la decidida colaboración del Ingeniero Roldán, se estudiaron las alternativas para aumentar la objetividad del proceso y conseguir un mayor cubrimiento en la muestra:

- Se decidió aumentar la sensibilidad del error estándar de 0.015 a 0.00016, con el fin de que este altere positivamente la varianza poblacional y proporcionalmente, aumente la muestra.
- Se decidió agrupar los departamentos en zonas geográficas (5), para clasificar la muestra y determinar así, la muestra por región y departamento.

## 2. ¿COMO SE REALIZA EL PROCESO?

**MATERIA PRIMA:** Para realizar el proceso, se necesita tener un archivo de Excel con todo el universo de planillas representadas con el código de la emisora, repetido  $n$  veces, con relación a la totalidad de planillas existentes para dicho código.

Este documento es generado o suministrado por la persona encargada de la recepción y codificación de las planillas de reporte del uso de la música:



	A	B	C	D	E
1	010110				
2	010105				
3	010112				
4	010110				
5	011001				
6	010303				
7	011001				
8	010109				
9	011001				

**FORMULACION:** Teniendo el universo ya determinado, se procede a definir las variables:

N: Totalidad de planillas.

$$n' = \frac{s^2}{\sigma^2} \begin{matrix} \longrightarrow & \text{Varianza muestral} \\ \longrightarrow & \text{Varianza Poblacional} \end{matrix}$$

$s^2$  : Se determina en términos de probabilidad como  $s^2 = p(1 - p)$  , siendo  $p=0.9$

$(se)^2$  : Se determina que este error estándar al cuadrado debe ser de 0.00016, y este, a su vez, determina la varianza poblacional.

Con los valores anteriormente mencionados, se determina la siguiente formula para la muestra:

$$n = \frac{n'}{1 + n'/N}$$

**CANTIDAD DE PLANILLAS A SELECCIONAR:** Se determina el Factor  $fh$  así:  
muestra seleccionada / universo:

$$fh = n / N$$

**AGRUPAMIENTO DE DEPARTAMENTOS EN REGIONES:** Se deben agrupar los departamentos en las 5 regiones geográficas:

Regiones				
<b>Id</b>	<b>CodRegion</b>	<b>REGION</b>	<b>Departamento</b>	<b>CodigoDepartamento</b>
33	1	AMAZONICA	Amazonas	32
34	1	AMAZONICA	Caquetá	14
35	1	AMAZONICA	Guaina	40
36	1	AMAZONICA	Guaviare	42
37	1	AMAZONICA	Putumayo	33
38	1	AMAZONICA	Vaupés	41
39	2	ANDINA	Antioquia	07
40	2	ANDINA	BOYACA	21
41	2	ANDINA	CALDAS	08
42	2	ANDINA	Cundinamarca	22
43	2	ANDINA	Huila	12
44	2	ANDINA	Norte de Santander	18
45	2	ANDINA	Quindío	10
46	2	ANDINA	Risaralda	09

Regiones				
Id	CodRegion	REGION	Departamento	CodigoDepartamento
47	2	ANDINA	Santander	19
48	2	ANDINA	Tolima	11
49	3	CARIBE	Atlántico	04
50	3	CARIBE	Bolívar	03
51	3	CARIBE	Cesar	20
52	3	CARIBE	Córdoba	01
53	3	CARIBE	Guajira	06
54	3	CARIBE	Magdalena	05
55	3	CARIBE	san andres	30
56	3	CARIBE	Sucre	02
57	4	ORINOQUIA	Arauca	31
58	4	ORINOQUIA	Casanare	44
59	4	ORINOQUIA	Meta	23
60	4	ORINOQUIA	Vichada	43
61	5	PACIFICA	Cauca	15
62	5	PACIFICA	Nariño	16
63	5	PACIFICA	Valle del Cauca	13
64	5	PACIFICA	CHOCO	17

**DIVISION DEL UNIVERSO:** por medio de los códigos que se encuentran en el listado inicial y que hacen referencia a la emisora-ciudad-departamento, se cruza la información, se clasifica y totaliza por regiones, así:

Muestra		
Id	CodRegion	Cantidad
6	1	129
7	2	1769
8	3	295
9	4	149
10	5	431

↓ Total del Universo **N**

**DETERMINANDO LA CANTIDAD DE PLANILLAS POR REGION:** Para determinar qué cantidad de planillas le corresponde a cada región, se procede multiplicando la cantidad de planillas por región por el factor obtenido anteriormente:

Muestra			
Id	CodRegion		Cantidad
6	1	Fn x	129
7	2	Fn x	1769
8	3	Fn x	295
9	4	Fn x	149
10	5	Fn x	431

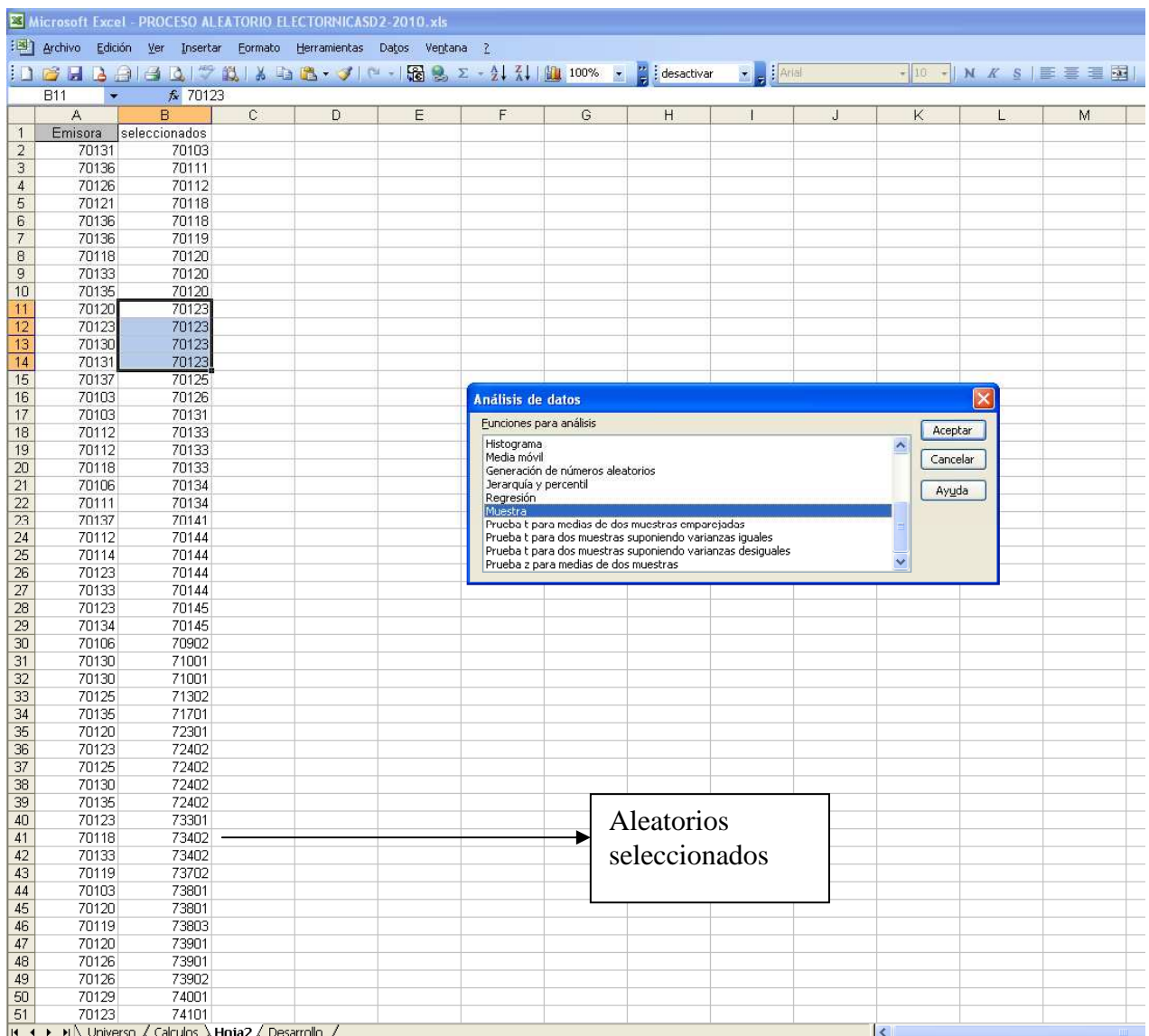
↓  
n1 +n2 +n3 +n4 +n5=n    Factor

El resultado de esta operación debe corresponder a la variable  $n$  (muestra) obtenida inicialmente.

Como ya se tiene la  $n$  por región, se determinan entonces de los  $X$  paquetes o planillas iniciales los  $n(i)$  aleatorios, así:

Se procede con la herramienta "Muestra" que se observa en la hoja de cálculo que se encuentra en la ruta: *herramientas/análisis de datos/muestra*.

En 2010 → Archivo → opciones → complementos → Administrar → complementos de Excel → ir → herramientas para análisis. Datos → análisis de datos → muestra.

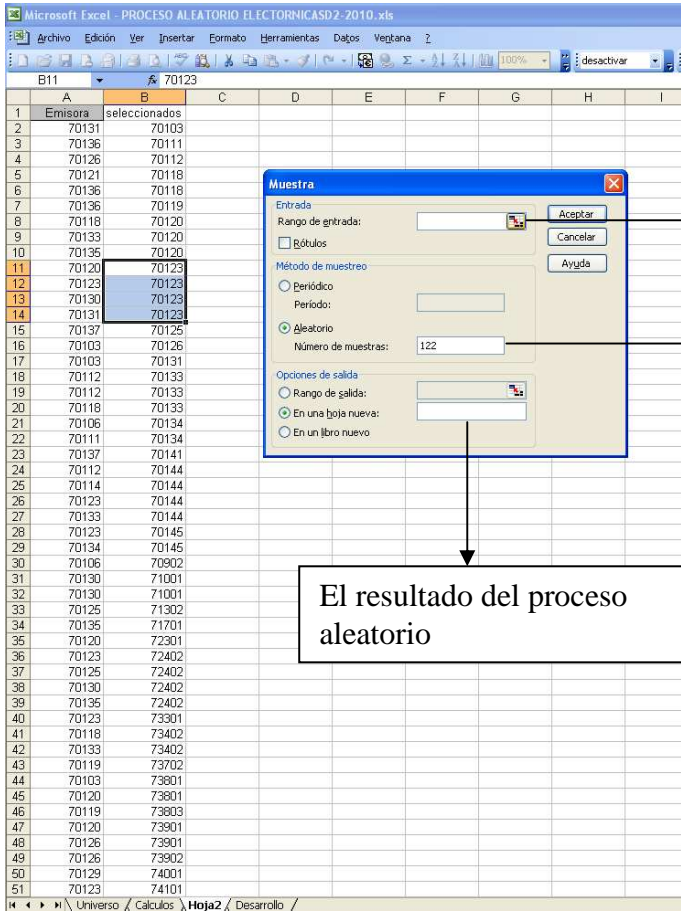


The screenshot shows an Excel spreadsheet with two columns: 'Emisora' (A) and 'seleccionados' (B). The data rows contain various alphanumeric codes. A blue box highlights a selection of rows in column B. An 'Análisis de datos' dialog box is open, with the 'Muestra' option selected under 'Funciones para análisis'. The dialog box includes options for 'Prueba t para medias de dos muestras emparejadas', 'Prueba t para dos muestras suponiendo varianzas iguales', 'Prueba t para dos muestras suponiendo varianzas desiguales', and 'Prueba z para medias de dos muestras'. Buttons for 'Aceptar', 'Cancelar', and 'Ayuda' are visible.

Emisora	seleccionados
70131	70103
70136	70111
70126	70112
70121	70118
70136	70118
70136	70119
70118	70120
70133	70120
70135	70120
70120	70123
70123	70123
70130	70123
70131	70123
70137	70125
70103	70126
70103	70131
70112	70133
70112	70133
70118	70133
70106	70134
70111	70134
70137	70141
70112	70144
70114	70144
70123	70144
70133	70144
70123	70145
70134	70145
70106	70902
70130	71001
70130	71001
70125	71302
70135	71701
70120	72301
70123	72402
70125	72402
70130	72402
70135	72402
70123	73301
70118	73402
70133	73402
70119	73702
70103	73801
70120	73801
70119	73803
70120	73901
70126	73901
70126	73902
70129	74001
70123	74101

Aleatorios  
seleccionados

↓  
Total paquetes de la región



Emisora	seleccionados
70131	70103
70136	70111
70126	70112
70121	70118
70136	70118
70136	70119
70118	70120
70133	70120
70135	70120
70120	70123
70123	70123
70130	70123
70131	70123
70137	70125
70103	70126
70103	70131
70112	70133
70112	70133
70118	70133
70106	70134
70111	70134
70137	70141
70112	70144
70114	70144
70123	70144
70133	70144
70123	70145
70134	70145
70106	70902
70130	71001
70130	71001
70125	71302
70135	71701
70120	72301
70123	72402
70125	72402
70130	72402
70135	72402
70123	73301
70118	73402
70133	73402
70119	73702
70103	73801
70120	73801
70119	73803
70120	73801
70126	73801
70126	75902
70129	74001
70123	74101

El origen son todos los paquetes de la región

Los  $n$  aleatorios que se deben seleccionar

El resultado del proceso aleatorio

Elaborado este proceso por región, se debe codificar entonces la totalidad de los paquetes y los seleccionados.

Region 1		
Emisora	CodRegion	seleccionado
140103	1	140202
140103	1	140107
140103	1	420107
140103	1	140107
140103	1	330601
140103	1	330601
140103	1	330601
140103	1	140106
140103	1	420103
140105	1	330601
140106	1	140106
140107	1	140111
140107	1	140801
140107	1	330501
140107	1	140103

planillas seleccionadas

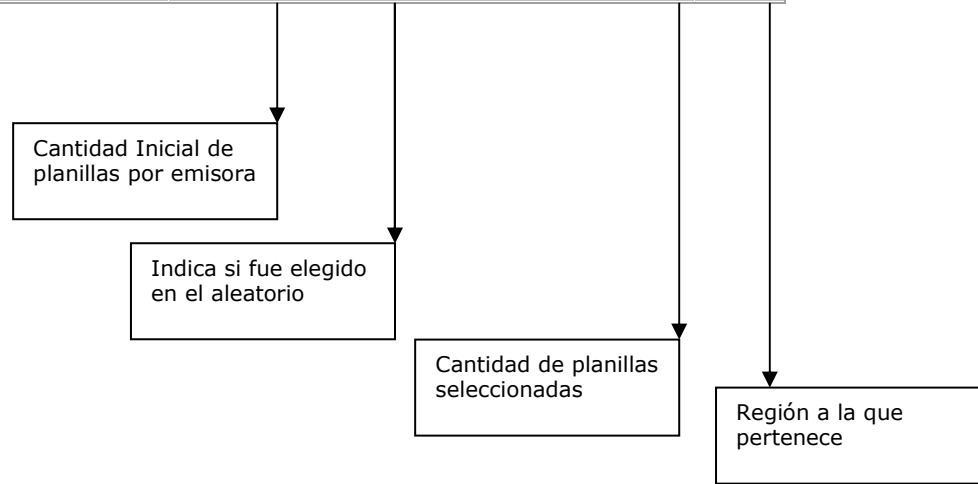
Region 1		
Emisora	CodRegion	seleccionado
140107	1	140111
140107	1	140111
140107	1	330501
140107	1	140107
140107	1	420107
140107	1	140801
140107	1	140105
140108	1	
140109	1	
140109	1	
140109	1	
140109	1	
140109	1	
140109	1	

Región

Terminada la codificación del aleatorio por región, se debe realizar la siguiente matriz de seguimiento:

EmisorasAgrupado						
Id	Codemisora	Emisora	Paquetes	Selección	Cantidad Seleccionada	Region
1	010101	EMISORA POLICIA NACIONAL	2	1	1	3
2	010105	FRECUENCIA BOLIVARIANA	2			3
3	010109	OLIMPICA FM ESTEREO	2			3
4	010110	RADIO TIEMPO FM ESTEREO	2			3
5	010112	TROPICANA	3	1	1	3
6	010201	RADIO CARACOL LORICA	1	1	1	3
7	010303	SION ESTEREO	10	1	2	3
8	010402	OMEGA STEREO	1	1	1	3
9	010502	MONTELIBANO FM STEREO	1			3
10	010901	BESAME SAN ANDRES DE SOTAVENTO	1			3
11	011001	EMISORA COLOMBIA STEREO 93 0	10	1	1	3
12	011201	CARIBE STEREO	3	1	1	3
13	020108	RADIO MINUTO	1			3
14	020110	OLIMPICA FM ESTEREO	2			3
15	020111	TROPICANA	3			3
16	030101	RADIO CIUDAD DE CARTAGENA COLMUNDO CARTAGENA	1			3
17	030105	RADIO BUCANERO	1	1	1	3
18	030106	RADIO ESPERANZA	11	1	4	3
19	030109	EMISORA MINUTO DE DIOS	9	1	2	3
20	030113	TIEMPO ESTEREO	5	1	1	3
21	030114	OLIMPICA FM ESTEREO	5			3
22	030116	EMISORA POLICIA NACIONAL	9	1	2	3
23	030117	RADIO RELOJ CARTAGENA	3	1	1	3

Emisoras Agrupado						
<b>Id</b>	<b>Codemisora</b>	<b>Emisora</b>	<b>Paquetes</b>	<b>Selección</b>	<b>Cantidad_ Seleccionada</b>	<b>Region</b>
24	030122	U DE C RADIO	6	1	1	3
25	030203	REIN FM ESTEREO	4			3
26	030301	GALAXIA STEREO	1			3
27	030601	TROPICANA CARTAGENA	4	1	2	3
28	030801	SIMITI ORIGINAL STEREO	1	1	1	3



### 3. CASOS CONTROLADOS:

Esta matriz permite determinar si los **paquetes seleccionados** exceden a los **paquetes iniciales**; esto ocurre, porque el aleatorio no es excluyente en su selección. En este caso, se debe realizar lo siguiente:

- Determinar cuales registros de la matriz son los que presentan esta característica
- Se debe determinar la diferencia, llevando al limite el total de paquetes seleccionados. Ejemplo:

*Emisora xxx: paquetes iniciales: 3 paquetes seleccionados: 5 diferencia: 2*

*Se debe entonces, llevar al máximo que es 3, y dejar los 2 como diferencia*

- La diferencia se debe distribuir como aleatorio entre los registros de la región, que no quedaron seleccionados.

Terminada la matriz se debe realizar el siguiente informe en formato PDF, para ser entregado al directo encargado de las planillas de reporte de uso de la música. Este informe debe contener la siguiente información: Subdivisión por Región-Departamento, emisora, cantidad de paquetes iniciales y cuantos paquetes fueron seleccionados para la muestra.



**DOCUMENTACION PROCESO  
ESTADISTICO INGRESO MUESTRA  
ELECTRONICA**  
Ingreso de Muestra  
ACINPRO  
2012

Adobe Reader - [INFORME.pdf]

Archivo Edición Ver Documento Herramientas Ventana Ayuda

Guardar una copia Buscar Seleccionar 110% Ayuda Buscar en Web Comparta fotografías como un profesional

**ACINPRO**  
**Planillas Seleccionadas D2 - 2010**

Region	Codemisora	Emisora	Paquetes	Selección
<b>AMAZONICA</b>				
<i>CAQUETA</i>				
	140103	ONDAS DEL ORTEGUAZA	9	1
	140105	CRISTALINA ESTEREO	1	1
	140106	ESPLENDIDA ESTEREO	1	1
	140107	EMISORA COLOMBIA MIA 910	11	3
	140111	EMISORA COLOMBIA STEREO 100 1	8	3
	140202	LINDA STEREO	5	1
	140801	COLOMBIA STEREO 103.9	13	2
	<b>Total Departamento</b>			<b>12</b>
<i>GUAVIARE</i>				
	420103	JUVENTUD ESTEREO	1	1
	420107	EMISORA POLICIA NACIONAL	6	2
	<b>Total Departamento</b>			<b>3</b>
<i>PUTUMAYO</i>				
	330203	LATINA STEREO	3	1
	330501	EMISORA COLOMBIA STEREO 94.7	13	2
	330601	OCAINA STEREO	8	4
	<b>Total Departamento</b>			<b>7</b>

Friday, September 10, 2010 Página 1 de 16

Archivos recientes

1 de 16

#### 4. ACTUALIZACIÓN ENERO DE 2012:

Detectada una falencia sobre la cantidad de tiempo transcurrida entre que una planilla es diligenciada por un radiodifusor y el momento en que es realmente utilizada para la muestra, se hace necesario plantear una estrategia que permita seleccionar de los paquetes aleatoriamente seleccionados y conforme la cantidad de planillas que indique dicho proceso, **las planillas cuya fecha de diligenciamiento, correspondan a las fechas más actualizadas, conforme la actual distribución.** Esta selección permite garantizar que la información ingresada al sistema se encuentra lo más actualizada posible, conforme la programación musical de las emisoras tenidas en cuenta en el proceso; con esto se busca tener una información más ajustada a la actualidad musical del país en el momento de la distribución correspondiente, sin afectar la aleatoriedad de la información.