

Redatam+SP GINIPlus

Cálculo del Índice de Gini Usando el REDATAM+SP

REDATAM+SP® es un programa computacional desarrollado por el Centro Latinoamericano y Caribeño de Demografía (CELADE), el cual es la División de Población de la Comisión Económica para la América Latina y el Caribe, (CEPAL), Naciones Unidas.

www.eclac.cl/celade/Redatam/

Tabla de Contenido

| | | |
|--------------|---|-----------|
| I. | Introducción..... | 2 |
| II. | Instalación..... | 2 |
| III. | Funcionamiento esquemático | 3 |
| IV. | Archivos INL de entrada..... | 3 |
| IV.1 | Archivo ALLINLFiles..... | 4 |
| IV.1.1 | Sección STRUCTURE | 4 |
| IV.1.2 | Sección INLFILES | 5 |
| IV.1.3 | Sección EXECINLS..... | 5 |
| IV.2 | Archivo INL de indicadores | 5 |
| IV.2.1 | Sección STRUCTURE | 6 |
| IV.2.2 | Sección INEQ..... | 6 |
| IV.2.3 | Sección SIDRA..... | 9 |
| V. | Interfaz de Entrada | 10 |
| V.1 | Panel General | 11 |
| V.2 | Panel Población Objetivo | 12 |
| V.3 | Panel Grupos Theil | 12 |
| V.4 | Panel Opciones SIDRA..... | 13 |
| V.5 | Panel Ejecuciones..... | 14 |
| V.6 | Barra de Comandos | 15 |
| VI. | Resultados | 16 |
| VI.1 | Resumen..... | 16 |
| VI.2 | Resultados detallados..... | 17 |
| VI.3 | Programa REDATAM ejecutado | 18 |
| VII. | Anexo 1 - Colores | 19 |
| VIII. | Anexo 2 – Memoria de Cálculo | 20 |
| VIII.1 | Curva de Lorenz y el Índice de Gini | 20 |
| VIII.2 | Índice de Theil..... | 21 |
| VIII.3 | Índice de Atkinson..... | 21 |
| VIII.4 | Línea de Pobreza | 21 |

I. Introducción

El objetivo de este documento es describir la aplicación GINIPlus para calcular el Índice de Gini de concentración (y otros índices correlatos, como el Theil y el Atkinson) usando el REDATAM+SP.

GINIPlus es un programa en los mismos moldes del xPlan, usando archivos de Indicadores de Lenguaje (INL) para ingresar los parámetros. La estrategia usada es construir un programa REDATAM en lenguaje SPC y procesarlo, y luego tomar las salidas de este programa y producir los indicadores.

El programa asume que las bases de datos REDATAM referenciadas en los archivos de parámetros estén instaladas y funcionando: el GINIPlus NO chequea la consistencia de las bases.

II. Instalación

Al ejecutar el programa SETUP y seguir sus instrucciones usando las opciones por defecto (Next... Next... Next...), será creado un directorio de instalación (C:\redatam\ginipius por defecto). Se sugiere que el nombre de este directorio no sea modificado para que los ejemplos instalados juntamente con el aplicativo funcionen perfectamente. Los siguientes archivos serán copiados al directorio:

| | |
|-------------------|--|
| GINIPLUS.exe | Programa ejecutable |
| GINIPLUS.tto | Opciones por defecto del programa |
| ALLINLFILES_*.inl | Archivos de partida del programa. Existe un archivo para cada idioma, solamente por cuestiones de documentación. Cuando ejecutado por la primera vez el programa pedirá el archivo de partida, y el usuario podrá elegir cualquier de ellos. |
| GINIPLUS*.wxd | Archivos de idiomas |
| GINI*.inl | Archivos de los Indicadores de Lenguaje (en el directorio EXAMPLE) |
| *.SEL | Archivos ejemplo de selecciones geográficas (en el directorio EXAMPLE) |
| BASE | Directorio de la base de datos ejemplo |
| GINIPLUS_ESP.PDF | Este documento |
| GINIPNAD | Directorio con todos los programas (INL) para el cálculo de los valores par ala publicación de la PNAD |

Si el GINIPlus es instalado en otro directorio que no sea el por defecto (c:\ginipius), será necesario editar el archivo ALLINFILES a ser usado para cambiar los nombres de los archivos INL mencionados, así como también editar estos archivos INL en el subdirectorio EXAMPLE para modificar las referencias a los archivos de selecciones geográficas (.SEL). Por último, abrir el diccionario de la base de ejemplo (subdirectorio BASE) en el programa Process de REDATAM y reapuntar sus archivos, guardando el diccionario al final.

El GINIPlus es auto suficiente en lo que se refiere al REDATAM+SP, es decir, él funciona independientemente, y no necesita que el REDATAM esté instalado.

III. Funcionamiento esquemático

GINIPlus es un aplicativo basado en la estrategia de generar un programa a medida de las necesidades del usuario, enviarlo a REDATAM, y procesar la salida intermedia de REDATAM, generando el resultado final, en este caso el Índice de Gini y otros indicadores correlatos.

La primera parte del programa (GINI Interfaz) empieza leyendo un archivo de control (ALLINL), el cual tiene una lista de posibles procesos a ser ejecutados (INL Proceso). Estos procesos pueden ser ejecutados de dos maneras: a) dinámicamente, es decir, a elección del usuario, llamándolo en pantalla; o b) automáticamente, todos de la lista.

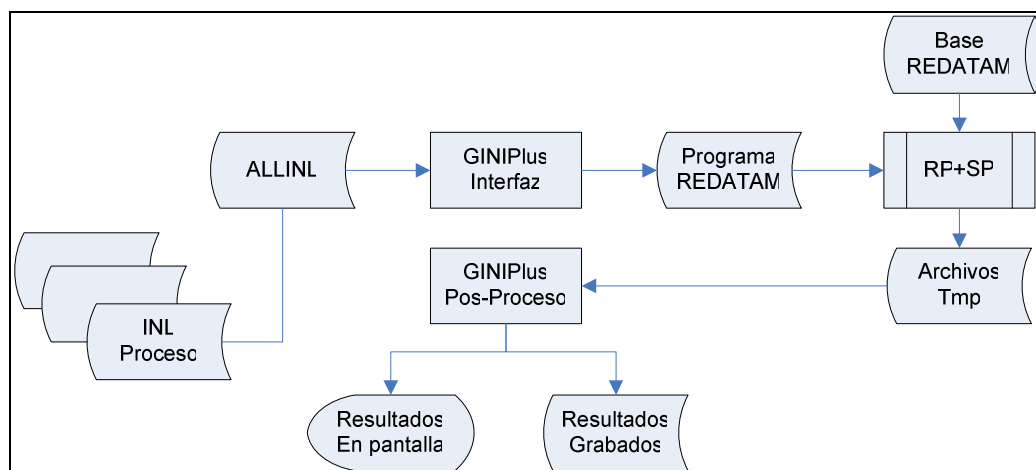


Figura 1 - Esquema de GINIPlus

Para cada proceso a ser ejecutado GINIPlus genera un programa en lenguaje interno de REDATAM (lenguaje pivot), el cual es ejecutado por el REDATAM. Este proceso es un conjunto de comandos DEFINE y dos tablas de AREALIST, una con una observación para cada elemento a ser contado, y otra con totales a nivel de la entidad de salida. Estos archivos son leídos de vuelta en el GINI (Pos-Proceso), y a partir de ellos se calculan los indicadores, los cuales son mostrados en pantalla y opcionalmente grabados en un archivo de salida.

IV. Archivos INL de entrada

Los archivos de entrada del GINIPlus poseen la extensión “INL”¹ de REDATAM, y se puede crear con cualquier editor de textos ASCII. Ellos son basados en una estructura de secciones y cláusulas, una en cada línea. Las secciones son informadas entre paréntesis cuadrados, por ejemplo,

```
[STRUCTURE]
```

Y las cláusulas son informadas con una parte de identificación (llave) y una parte de valor, separadas por una señal de igual (“=”), sin espacios entre la llave y la señal, y entre la señal y el valor. Estos valores para las cláusulas también son llamados de parámetros.² Ejemplo,

```
LLAVE=valor
```

¹ Del nombre, en inglés, Indicator Language, para la construcción de indicadores

² A lo largo de este documento, llamamos de una “entrada” a una línea con el conjunto de “cláusula=parámetro”

Las secciones no tienen un orden específico en el archivo, así como las cláusulas tampoco tienen un orden específico dentro de la sección. Líneas en blanco pueden ser usadas para dar más claridad y separar las secciones y cláusulas. Toda línea que no sea reconocida como una sección o una llave es considerada comentario. Para efectos de estandarización se usa, en general, un asterisco (“*”) o dos barras (“//”) en el inicio de una línea de comentario. Ej:

```
*****
* Declaración del diccionario (obligatorio)
*****
```

IV.1 Archivo ALLINLFiles

Este archivo es el que comanda inicialmente el GINIPlus. Su función es informar dos listas de archivos de indicadores que serán ejecutados por el GINIPlus, una lista de ejecuciones dinámicas y otra lista de ejecuciones automáticas. El archivo tiene tres secciones, todas opcionales. Para mayores detalles, ver el archivo ALLINLFiles_Esp.INL de ejemplo.

La diferencia entre las ejecuciones dinámica y automática es que en la primera el usuario puede mirar los parámetros informados, cambiarlos, si necesario, y luego ejecutar el Indicador. Es decir, los indicadores tienen que ser ejecutados uno a uno. La ventaja de este método es que los parámetros no tienen que estar todos informados *a priori*, ellos pueden ser modificados por el usuario antes de cada ejecución.

Por otro lado, la ejecución automática es una especie de ejecución en “batch”, todos los archivos en la lista son ejecutados en secuencia, uno tras el otro, sin intervención del usuario. Por ese motivo, todos los parámetros de cada indicador tienen que estar informados, así como sus valores iniciales, para que el programa pueda ejecutarlos automáticamente. La ventaja de ese método es que se puede pre-programar una lista de indicadores y ejecutarla de una sola vez.

| | | |
|-----------|----------|--|
| STRUCTURE | opcional | Título de la forma e idioma de trabajo |
| INLFILES | opcional | Archivos de ejecución dinámica |
| EXECINLS | opcional | Archivos de ejecución automática |

Las secciones INLFILES y EXECINLS son opcionales, pero una de las dos tiene que estar presente.

IV.1.1 Sección STRUCTURE

Esta sección tiene dos cláusulas, FORMCAPTION y LANGUAGE.

| Cláusula | Comentario |
|-------------|---|
| [STRUCTURE] | |
| FORMCAPTION | Título de la forma |
| LANGUAGE | Idioma de trabajo, necesario solo si es distinto del informado en el archivo TTO Esp – para español Prt – para portugués Eng – para inglés Fra – para francés |

Ejemplo:

```
FORMCAPTION=Indice de GINI y otros Indicadores
LANGUAGE=Esp
```

IV.1.2 Sección INLFILES

Esta sección tiene tres cláusulas, FILES, LAST y FILENAMEi. Ella informa los archivos de indicadores que podrán ser ejecutados dinámicamente, bajo comando directo del usuario.

| Cláusula | Comentario |
|-------------------|--|
| [INLFILES] | |
| FILES | Número de archivos INL que vienen en las cláusulas FILENAME |
| LAST | Número del archivo que será mostrado al abrir el programa. Opcional, si no existe muestra el primero |
| FILENAMEi | Nombre completo (nombre y paso) del archivo INL de indicadores |

Ejemplo:

```
FILES=6
LAST=2
```

```
FILENAME1=C:\GINIPLUS\example\GINI_TOTAL_Esp.inl
FILENAME4=C:\GINIPLUS\example\GINI_REG_T.inl
FILENAME5=C:\GINIPLUS\example\GINI_BR_H.inl
```

No es necesario que vengan todos los FILENAMEi, como se ve en el ejemplo, no vino el FILENAME2.

IV.1.3 Sección EXECINLS

Esta sección tiene también tres cláusulas, FILES, FILENAMEi y FILETYPEi. Ella informa los archivos de indicadores que serán ser ejecutados automáticamente.

| Cláusula | Comentario |
|-------------------|---|
| [EXECINLS] | |
| FILES | Número de archivos INL que vienen en las cláusulas FILENAME |
| FILENAMEi | Nombre completo (nombre y paso) del archivo INL de indicadores |
| FILETYPEi | Tipo de ejecución (ALL GINI SIDRA). Opcional, si no existe es ALL |

Ejemplo:

```
FILES=5
```

```
FILENAME1=C:\GINIPLUS\example\GINI_REG_T.inl
FILENAME3=C:\GINIPLUS\example\GINI_2_T.inl
FILETYPE3=SIDRA
FILENAME4=C:\GINIPLUS\example\GINI_BR_H.inl
FILETYPE4=GINI
```

No es necesario que vengan todos los FILENAMEi, como se ve en el ejemplo, no vino el FILENAME2.

IV.2 Archivo INL de indicadores

Estos son los que verdaderamente tienen los parámetros de ejecución del GINIPlus. El archivo tiene tres secciones.

| | | |
|-----------|-------------|---------------------------------------|
| STRUCTURE | obligatoria | Título del indicador |
| INEQ | opcional | Parámetros de cálculo del Gini |
| SIDRA | opcional | Parámetros de grabación para el SIDRA |

La más importante de ellas es la [INEQ], la cual trae los parámetros de cálculo del Gini. En principio esta sección debiera ser obligatoria, y casi siempre así lo es, pero hay casos en que se necesita ejecutar solamente el SIDRA, luego de una ejecución del GINI, y en ese caso basta informar los parámetros SIDRA (ver V.4).

Algunas cláusulas son obligatorias, otras son opcionales, y otras son más que opcionales, es decir, innecesarias. Es el caso de las cláusulas que tienen la palabra CAPTION como parte del nombre, y son usadas solamente para modificar los textos de las pantallas. Estos textos son establecidos por el archivo de multi idiomas (GINIPlusEsp.WXD), y tendrían que ser modificados solamente en el caso de que estuvieran claramente inapropiados. Por ejemplo, cuando se calcula el GINI para otro tipo de dato que no sea el de ingreso o de sueldo, se podría cambiar uno u otro título para corresponder a los datos usados. En el ejemplo, todas esas cláusulas están separadas en el archivo TRANSLATE_Esp.INL, el cual puede (o no) ser usado en el archivo principal de indicadores con un comando #INCLUDE. De existir, ese comando solo puede estar en la sección [INEQ].

IV.2.1 Sección STRUCTURE

Esta sección es usada para informar parámetros generales del programa. La cláusula más importante e imprescindible es la del WORKSPACE, la demás son opcionales.

| Cláusula | Comentario |
|-------------|--|
| [STRUCTURE] | |
| CAPTION | Título para la caja de índices a ser ejecutados (opcional). Por defecto es el nombre del archivo INL |
| WORKSPACE | Declaración del nombre completo del diccionario de la base de datos (obligatorio) |
| COLOR | Color de la línea diagonal de la curva de Lorenz. Acepta un nombre de color (ver VII. Anexo 1 - Colores) o un número en formato RGB. De no ser informada asume ROJO |
| COLORMIN | Color inicial para las curvas de Lorenz cuando existir más de un área geográfica para representar, o cuando exista solamente una curva. Por defecto es BLANCO |
| COLORMAX | Color final para las curvas de Lorenz cuando existir más de un área geográfica para representar. Por defecto es AZUL |
| GRAPHMODE | Tipo de gráfico para la Curva de Lorenz. Actúa en el límite máximo del eje vertical. Si EXTENDED, el límite es el promedio de los valores, y si STANDARD, el límite es 100 (los valores de la curva son transformados para el intervalo 0-100). Por defecto asume STANDARD. ³ |

Ejemplo:

```
WORKSPACE=C:\GINIPLUS\BASE\PNAD_PARTE.dic
CAPTION=Ejemplo completo
COLOR=128.0.0
COLORMIN=BLUE
COLORMAX=GREEN
GRAPHMODE=EXTENDED
```

IV.2.2 Sección INEQ

Esta sección es usada para llenar los tres paneles principales del GINIPlus, el panel General, el panel de Selección de la Población Objetivo, y el panel de los Grupos Theil. Ella tiene varios bloques, uno para cada uno de los controles de los paneles mencionados. En general hay dos tipos de controles: a) una lista de posibilidades; y b) un valor único. En el caso del valor único la cláusula es sencilla, con una sola información, pero para las listas, hay que tener tres datos: 1) el número

³ El STANDARD es especialmente ventajoso cuando hay que desplegar más de una curva, porque se convierten todas ellas para el mismo intervalo

de elementos de la lista; 2) los elementos propiamente tales; y 3) cual es el elemento por defecto (el elemento inicial, el que aparece en la lista).

Controles de varias posibilidades

Las cláusulas son construidas con el nombre del CONTROL y las siguientes terminaciones:

| Terminación | Comentario |
|-------------|---|
| N | Número de elementos. Ej: SELECTIONN=5 hay 5 elementos para el control SELECTION (selecciones geográficas) en la lista. Los elementos no pueden tener fallas en la secuencia |
| i | i es un número, y representa un secuencial, para los elementos del control. Ej: GINIMBR2=DOMICIL.SETORSIT el elemento 2 en la lista GINIMBR (variables para el cálculo del GINI) es la variable SETORSIT de la entidad DOMICIL |
| LAST | Número del elemento por defecto a aparecer en el control. Ej: THEILVARLAST=4 El elemento por defecto del control THEILVAR (variables para filtros en THEIL) es el 4 |
| i.VLj | Cuando los elementos de una lista son variables de la base, y se quiere referenciar a la categoría j de la variable i en la lista. Ej: GINIMBR1.VL2=YES la categoría 2 de la variable 1 en la lista para el control GINIMBR (variables para el cálculo del GINI) va a estar seleccionada (pré-marcada) |

| Control | Comentario |
|------------------|--|
| SELECTION | Selecciones geográficas disponibles (opcional). De no existir o SELECTIONN=0, el control no es mostrado en pantalla y asume ALL. Si al ejecutar, el archivo informado no existir, asume ALL. Los elementos de la lista son nombres completos (paso y nombre) de archivos de selecciones geográficas (con terminaciones .sel o .slw). Ej: SELECTION2=C:\GINIPLUS\example\AMAPA.SEL Tiene también la cláusula SELECTIONFIXED para informar una sola selección. En ese caso la selección única no es desplegada en pantalla. De existir, esa cláusula tiene precedencia sobre el número de selecciones. Ej: SELECTIONFIXED=SELECTION4=C:\GINIPLUS\example\RNORTE.SEL Excepcionalmente este control acepta fallas en la secuencia de sus elementos |
| GINIVALUE | Variables de la base para el cálculo del Gini y otros indicadores. Obligatoria la existencia de por lo menos una variable en la lista. Deben ser variables cuantitativas, tipo Ingreso de la persona (o del hogar), o área del establecimiento agropecuario. Los elementos de la lista son nombres completos de variables (entidad.variable). Ej: GINIVALUE4=PESSOA.V4720 |
| CNT | Entidades para elegir cual de ellas será contabilizada. Obligatoria la existencia de por lo menos una entidad en la lista. Los elementos de la lista son entidades reales (no pueden ser virtuales) de la base. Ej: CNT3=PESSOA |
| WEIGHT | Peso a ser usado en el proceso (opcional). De no existir, o WEIGHTN=0 el control no es mostrado en pantalla, y no se usará el peso en el cálculo, a no ser que la base tenga peso automático en las entidades, y no sea informada la cláusula WEIGHTAUTO=NO. Los elementos de la lista son variables |

| | |
|-----------------|---|
| | de la base, o un texto cualquier, entre paréntesis, para significar que no se usará el elemento. Ej: WEIGHT1=(Sin peso) Hay también la cláusula WEIGHTAUTO (por defecto es "NO") para usar el peso automático si la base tiene pesos automáticos para las entidades. La cláusula WEIGHTAUTO tiene precedencia sobre WEIGHTN, es decir, si ambas están informadas, el sistema usa el WEIGHTAUTO y desconsidera el WEIGHTN |
| GINIOUT | Nivel de salida. Obligatoria la existencia de por lo menos un nivel, aunque sea la raíz (toda la encuesta). Los elementos de la lista son entidades SELECCIONABLES de la base. Ej: GINIOUT2=REGIAO |
| GINIMBR | Filtro adicional (opcional) para el GINI. De existir, este filtro será combinado (AND) con el filtro de expresión (FILTER). Los elementos de la lista son nombres completos de variables CATEGORIZADAS de la base. Acepta la pré-marcación de categorías. Ej: GINIMBR3=DOMICIL.V0202 GINIMBR1.VL2=YES |
| POINT | Número de puntos en la Curva de Lorenz (opcional). De no existir o POINTN=0 asume 2, 4, 5, 10, 20 y 100 puntos. Los elementos de la lista son expresiones con un número (valor máximo = 1000) seguido de un texto. Si no es un número el elemento es descartado. Ej: POINT2=50 puntos Tiene también la cláusula POINTFIXED para informar una sola posibilidad de puntos. En ese caso la caja de posibilidades no es desplegada en pantalla. De existir, esta cláusula POINTFIXED tiene precedencia sobre el número de puntos. Ej: POINTFIXED=20 El elemento por defecto (POINTLAST) es una excepción a la regla general de informar el número del elemento. En este caso tiene que informar el número de puntos. Ej: POINTLAST=100 Excepcionalmente, este control acepta fallas en la secuencia de sus elementos |
| THEILVAR | Parámetros para los filtros de Theil (opcional). Los elementos de la lista son nombres completos de variables CATEGORIZADAS de la base. Acepta la pré-marcación de categorías. Ej: THEILVAR4=DOMICIL.V0202 THEILVAR2.VL1=YES |

El control "FILTER" es la excepción para las reglas de nombramiento de las cláusulas, por eso es mostrado a parte.

| Cláusula | Comentario |
|-----------------------|---|
| FILTERFIXED | Expresión de filtro, única. Ej: FILTERFIXED=PESSOA.IDADE >= 10 Esta cláusula tiene precedencia sobre FILTERN. |
| FILTERN | Número de elementos en la lista. |
| FILTERVALUEi | Expresión del filtro en lenguaje SPC. Ej: FILTERVALUE1=DOMICIL.ESPECIE < 5 Excepcionalmente ese control acepta fallas en la secuencia de sus elementos |
| FILTERCAPTIONi | Texto descriptivo para el elemento del filtro en la caja (opcional). De no informar se usa la propia expresión del filtro. Ej: Viviendas particulares permanentes |

Controles de valor único

| Cláusula | Comentario |
|----------------|---|
| AREALABELCHECK | Nombres de área para los elementos de la entidad de salida (opcional). Por defecto es "YES", siempre y cuando la entidad de salida tenga nombres de áreas. Ej: AREALABELCHECK=YES |
| EXTRAIDENTLAST | Identificación extra para la salida (opcional). Caso la identificación de la entidad de salida no sea el código compuesto, elegir cual la variable de la lista debe ser usada (en el caso de bases como la PNAD, que tiene una entidad con el año de la encuesta, este campo es necesario para identificar solamente la REGIAO y UF, por ejemplo, sin agregar el ANO de la encuesta en el código de identificación). El GINIPlus arma automáticamente la lista de las posibles identificaciones con todas las variables de tipo STRING de la entidad de salida. EXTRAIDENTLAST=3 |
| POVERTYLINE | Valor de ingreso abajo del cual las personas son consideradas pobres (opcional). Ej: POVERTYLINE=500 funciona como un filtro adicional para considerar solamente los casos en los cuales el valor de la variable objetivo es MENOR que el valor informado en la cláusula |
| MISSINGVALUE | Valor que simula el inválido en la base (opcional). Ej: MISSINGVALUE=999999999 funciona como un filtro adicional para considerar solamente los casos en los cuales el valor de la variable objetivo es DIFERENTE que el valor informado en la cláusula |
| ZEROVALUECHECK | Omitir valores con cero (opcional). Por defecto es "NO", es decir, considerar los casos con valor cero. Ej: ZEROVALUECHECK=YES funciona como un filtro adicional para considerar solamente los casos en los cuales el valor de la variable objetivo es DIFERENTE de cero |
| OUTPUTFILE | Nombre completo (paso y nombre) para el archivo de salida para los resultados del GINI (opcional). Ej: OUTPUTFILE=C:\GINIPLUS\EXAMPLE\Output\gini1_OUTPUT.txt |

IV.2.3 Sección SIDRA

Esta sección es usada para definir cuales valores calculados por el GINI se debe pasar al SIDRA (ver panel de Opciones SIDRA). Recuerde que el GINI calcula una serie de valores, como el propio GINI, así como los promedios, promedios acumulados, porcentajes y porcentajes acumulados. Además, hay que informar también algunos parámetros propios del SIDRA, tales como la identificación de la tabla, el año de la encuesta, etc.

| Cláusula | Tipo | Comentario |
|----------------|---------|--|
| [SIDRA] | | |
| SIDRAOPTION | integer | Elegir cual el valor a ser pasado para el SIDRA * 1 - GINI * 2 - Valor Promedio * 3 - Promedio Acumulado * 4 - Porcentaje * 5 - Porcentaje Acumulada |
| SIDRAPARM | texto | Conjunto de parámetros específicos del SIDRA, tales como tabla, año, variable, etc. Para más detalles, ver el documento "Usando o REDATAM para cargar dados agregados na base SIDRA" |

| | | |
|-------------|---------|---|
| SIDRAOUTPUT | texto | Nombre completo (nombre y paso) del archivo de salida del SIDRA |
| SIDRASTART | integer | Punto de partida para comenzar a generar resultados para el SIDRA. De no informar asume 1. El 0 es un valor especial para señalar solamente el valor total. |
| SIDRAVALUE | integer | Valor de partida para recodificar las categorías para el SIDRA. De no informar asume 1 |

Ejemplo:

```
[INEQ]
POINTSLAST=20
[SIDRA]
SIDRAOPTION=2
SIDRAPARM=T=1975 A=2005 V=4678 D=5 N=1 S1=5
SIDRAOUTPUT=C:\GINIPLUS\EXAMPLE\OUTPUT\T1975_gini1.txt
SIDRASTART=20
SIDRAVALUE=12
```

El control SIDRAOPTION=2 es para tomar el valor promedio. Si el número de puntos en la ejecución es 20, el SIDRASTART=20 significa que solamente el último valor será pasado al SIDRA (con 20 puntos eso significa que será el valor del 95%), y ese valor tendrá la categoría de 12 informada por el STARTVALUE.

Las cláusulas SIDRASTART y SIDRAVALUE no son usadas con el SIDRAOPTION=1 (GINI), porque esta cláusula pasa solamente un valor al SIDRA. Su utilización está más bien definida en V.4.

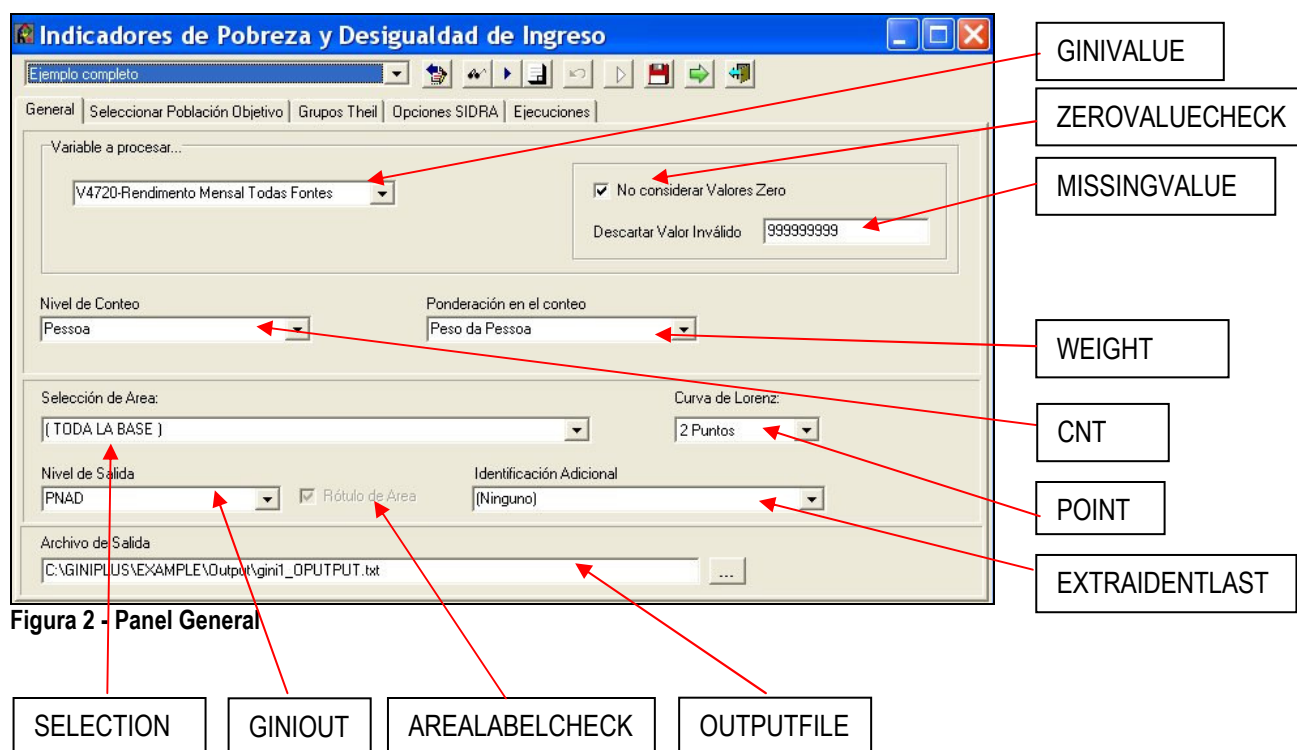
V. Interfaz de Entrada

GINIPlus tiene 5 paneles de entrada, mostrados en las secciones a seguir. En esas secciones se muestran las figuras de cada panel acompañadas de cajas con los nombres de los controles a ser usados, y flechas indicando el local de cada control. Además, las secciones traen informaciones adicionales sobre los controles usados. La ejecución del programa es totalmente controlada por la Barra de Herramientas, mostrada en la sección V.6.

Los paneles de entrada son mostrados solamente si hay controles informados para ellos. El panel General siempre aparece porque tiene controles obligatorios.

Pasando el Mouse sobre los controles (y sobre los botones de la Barra de Comandos) el sistema muestra una pequeña descripción de su funcionalidad.

V.1 Panel General



El control POINT es usado para definir el número de puntos en la Curva de Lorenz, lo que puede parecer una necesidad “cosmética”, en consecuencia también define, en el resultado detallado, el número de líneas de valores para el cálculo de los porcentajes. Por ejemplo, si el número de puntos elegido es 2, la Curva de Lorenz se resumirá a dos segmentos de recta, lo que no tiene mucha representatividad gráfica, pero en el resultado detallado será mostrado el valor de la mediana. Si el número de puntos es 10, tendremos los deciles, y si es 100, los percentiles. Ese control **NO AFECTA PARA NADA** el cálculo del Índice de Gini.

El control EXTRAIDENTLAST es muy específico, solo es usado en encuestas con una identificación de casos en la cual se tiene el periodo de la encuesta, por ejemplo, además de la identificación geográfica.

Los controles AREALABELCHECK y OUTPUTFILE tampoco son necesarios para la ejecución del programa. El primero despliega el nombre de las áreas en los resultados detallados, y el segundo sirve para grabar los resultados en un archivo de salida, luego de la ejecución del cálculo, en la ejecución dinámica (a elección del usuario). En la ejecución automática, un archivo INL sin el control OUTPUTFILE no tendrá sus resultados grabados.

El archivo definido en OUTPUTFILE será grabado en formato texto (ASCII). El botón al lado del control OUTPUTFILE es para navegar en el disco y elegir el nombre del archivo de salida.

V.2 Panel Población Objetivo

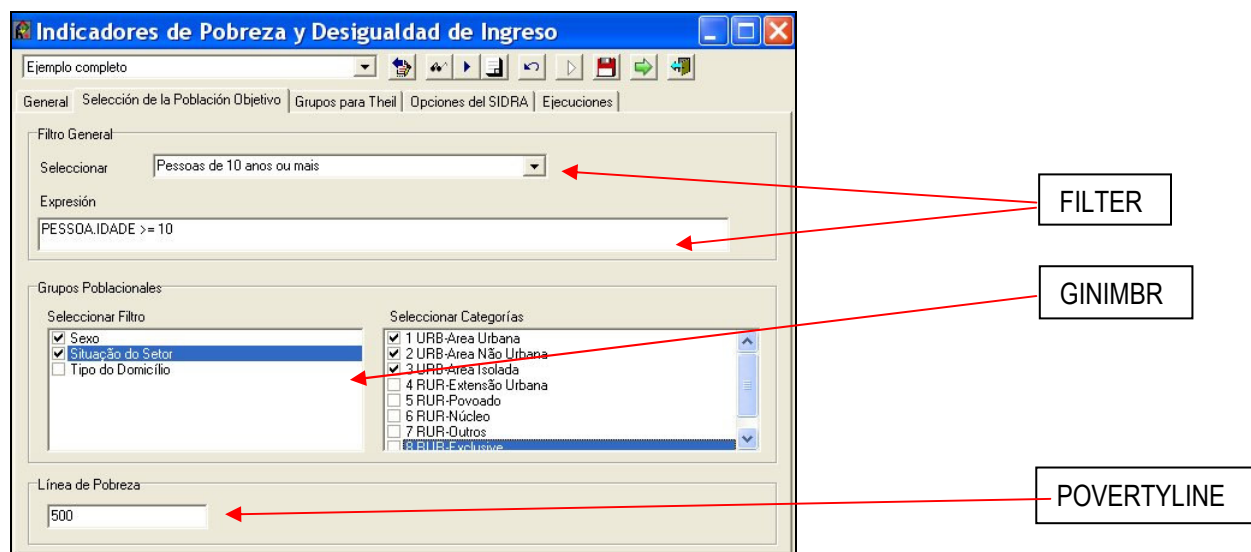


Figura 3 - Panel de Selección

Los controles FILTER y GINIMBR son usados en una combinación lógica (AND). El FILTER es más usado para variables cuantitativas, mientras que el GINIMBR es solamente para variables categorizadas. En la ventana del GINIMBR, todas las variables marcadas son usadas como filtro, aunque sus categorías no aparezcan en pantalla.

El control POVERTYLINE, además de forzar el cálculo para un determinado grupo de casos (los que estén por debajo de la línea de pobreza definida por el valor del parámetro), también arroja a la salida resultados específicos, como el Índice de Sen.

V.3 Panel Grupos Theil

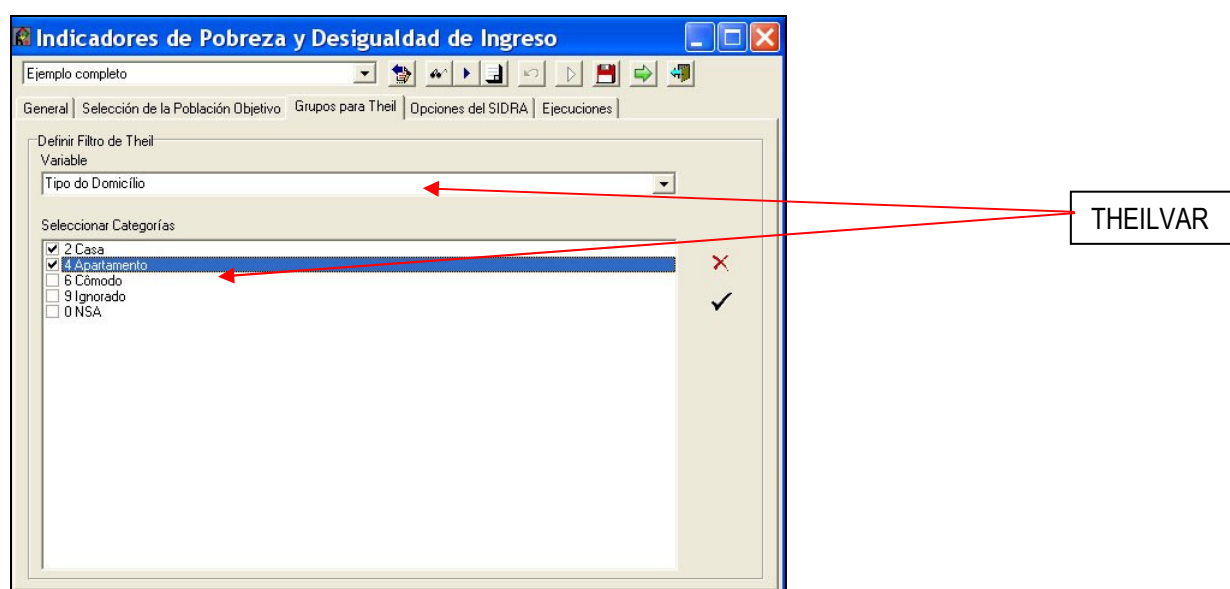


Figura 4 - Panel de los Grupos de Theil

El control THEILVAR, si usado, además de los índices de THEIL y THEIL DUAL (siempre calculados por defecto), saca los resultados para los grupos de categorías definidos, la Redundancia, etc.

V.4 Panel Opciones SIDRA

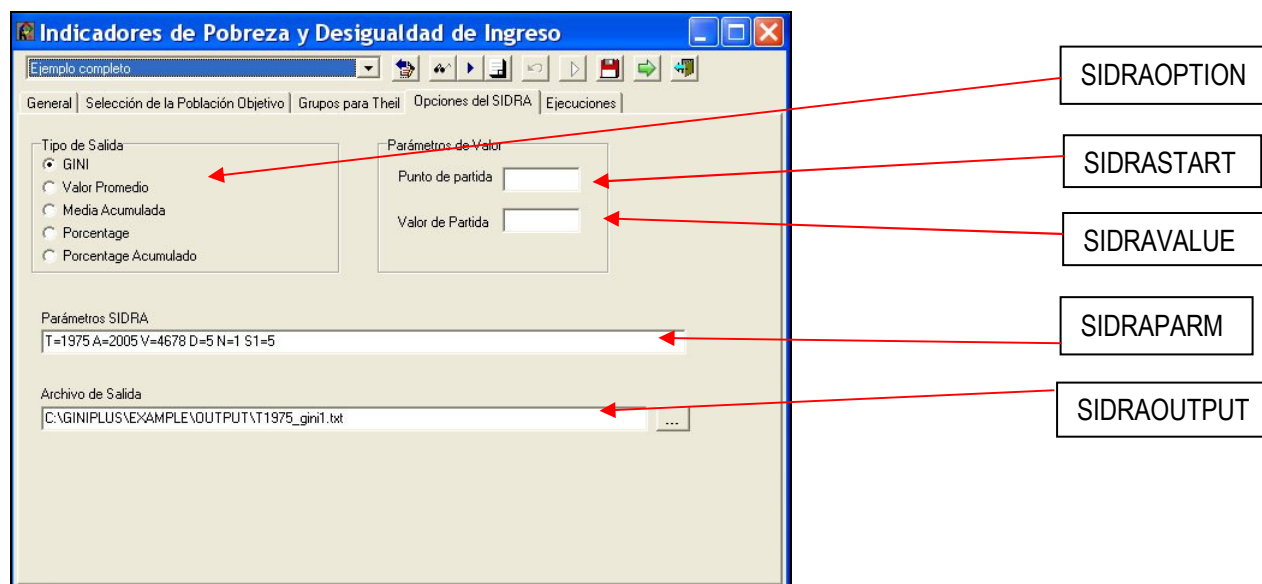


Figura 5 - Panel de Opciones SIDRA

Los controles SIDRASTART y SIDRAVALUE son usados para determinar cuales valores de salida serán pasados al SIDRA (SIDRASTART), y con cuales códigos de identificación (SIDRAVALUE). El número de valores existentes y disponibles es determinado por el número de puntos de la Curva de Lorenz (ver control POINT). Si el número de puntos es 10, por ejemplo, y no se informa nada en SIDRASTART, los 10 valores serán pasados al SIDRA. Sin embargo, si SIDRASTART=8, solamente los últimos 3 valores (8, 9 y 10) serán pasados al SIDRA.

El control SIDRAVALUE es para cambiar la identificación por defecto de cada valor (secuencial, empezando de 1). Para el ejemplo anterior (de 3 valores pasados al SIDRA), de no informarse el SIDRAVALUE, los valores serán identificados (las categorías) por 8, 9 y 10. Pero si SIDRAVALUE=15, los valores serán identificados por las categorías 15, 16 y 17.

Un ejemplo práctico de esos controles es la siguiente tabla de promedios, que tiene el promedio total, los promedios de 10 en 10%, el promedio del 95% y el promedio del 99%.

| Código | Contenido |
|--------|-----------------|
| 1 | Total |
| 2 | Hasta 10 |
| 3 | Más de 10 a 20 |
| 4 | Más de 20 a 30 |
| 5 | Más de 30 a 40 |
| 6 | Más de 40 a 50 |
| 7 | Más de 50 a 60 |
| 8 | Más de 60 a 70 |
| 9 | Más de 70 a 80 |
| 10 | Más de 80 a 90 |
| 11 | Más de 90 a 100 |
| 12 | Más de 95 a 100 |

Esa tabla necesita de 4 ejecuciones distintas del GINIPlus: a) para el total; b) para los resultados deciles; c) para el resultado del 95%; y d) para el 99%. La columna del “Código” equivale al valor del parámetro SIDRAVALUE, o sea, es la categoría de la variable que el SIDRA espera recibir junto con los valores, y poder identificarla y poner los valores en las celdas apropiadas.

Los controles serían:

```
*1ra ejecución GINI
POINTLAST=10
*para el total
SIDRASTART=0
SIDRAVALUE=1
*para los deciles
SIDRASTART=1
SIDRAVALUE=2
*2da ejecución GINI
POINTLAST=20
SIDRASTART=20
SIDRAVALUE=12
*3ra ejecución GINI
POINTLAST=100
SIDRASTART=100
SIDRAVALUE=13
```

Sería posible aumentar la eficiencia del proceso combinando las dos ejecuciones iniciales en una única ejecución del cálculo de los valores, pero con 2 salidas al SIDRA, una para el total y otra para los deciles).

El botón al lado del control SIDRAOUTPUT es para navegar en el disco y elegir el nombre del archivo de salida.

V.5 Panel Ejecuciones

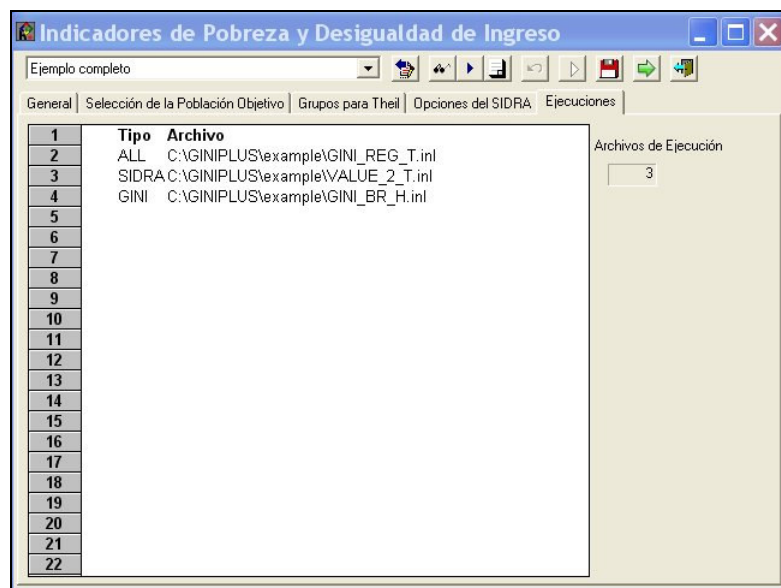


Figura 6 - Panel de Ejecuciones

Este panel no tiene ningún control operado por el usuario. Su utilidad es solamente para informar cuales archivos serán ejecutados automáticamente al apretar el botón de ejecución automática (ver V.6). Durante la ejecución una marca aparecerá al lado del nombre de cada archivo para informar el andamiento del proceso.

V.6 Barra de Comandos



Figura 7 - Barra de Comandos

Caja de INLs Recarga Compilar Ejecutar Mostrar salidas Tipo de Gráfico Ejecución automática Guardar resultados Exportar al SIDRA

Pasando el Mouse sobre el control el sistema muestra una pequeña descripción de su función.

| | |
|---------------------------|--|
| Caja de INLs | Para elegir el archivo INL a ser ejecutado por el botón de Ejecutar. Los controles de ese archivo serán cargados en pantalla, y podrán tener sus valores modificados antes de la Ejecución. |
| Recarga | Es una facilidad operacional caso sea necesario hacer algún cambio en el contenido del archivo INL en memoria (apuntado por la Caja de INLs), usando cualquier editor de texto. Luego de esos cambios, y de guardar el archivo en el editor, apretar ese botón para hacer la recarga del INL. ⁴ |
| Compilar | Este botón muestra el programa que será enviado al Process de REDATAM para ser ejecutado y luego tener sus resultados leídos por el GINIPlus en la segunda etapa (Pós-Proceso). El programa es mostrado en lenguaje SPC, tal como se aprecia en VI.3. |
| Ejecutar | Ejecuta el archivo INL en memoria. |
| Mostrar salidas | Es un botón para alternar la pantalla del programa entre los paneles de salida y los paneles de entrada. |
| Tipo de Gráfico | Alterna entre los tipos de gráfico STANDARD y EXTENDED. La diferencia básica entre los dos es que el EXTENDED usa como límite superior el valor promedio de la variable, y el STANDARD usa como límite superior el número 100, haciendo una razón entre el mayor de los valores máximos y el 100. El tipo EXTENDED debe ser más usado para mostrar resultados de una única curva, mientras que el STANDARD sirve más para mostrar resultados de varias curvas. |
| Exportar al SIDRA | Toma el resultado de los cálculos del GINIPlus y, usando los controles definidos en el panel SIDRA, exporta el resultado elegido para el archivo de salida en formato SIDRA definido en el control SIDRAOUTPUT. |
| Guardar resultados | Toma el resultado de los cálculos del GINIPlus y los exporta a un archivo de tipo texto definido en el control OUTUTFILE. |

⁴ El botón recarga solamente el INL en memoria. Para recargar la lista de INLs hay que reejecutar el programa, cerrándolo y abriéndolo de nuevo.

Ejecución automática Ejecuta en secuencia todos los archivos INL informados en la lista de EXECINLs. Es útil cuando se tiene un conjunto de cálculos a ser ejecutados para un plan tabular. Se programan los archivos INL de cada uno de los cálculos, se arma la lista en el archivo ALLINLFiles, se ejecuta el GINIPlus para cargar la lista en memoria, y entonces se usa ese botón para ejecutar la lista.

VI. Resultados

Los resultados son mostrados en tres diferentes paneles: a) Resumen; b) Detalle; y c) Programa compilado. La elección de cual panel es mostrado se hace en la caja combinada de Tipo de Salida, la cual tiene las posibilidades de Resumen, una línea para cada área de salida, y la opción para ver el Programa Compilado.

VI.1 Resumen

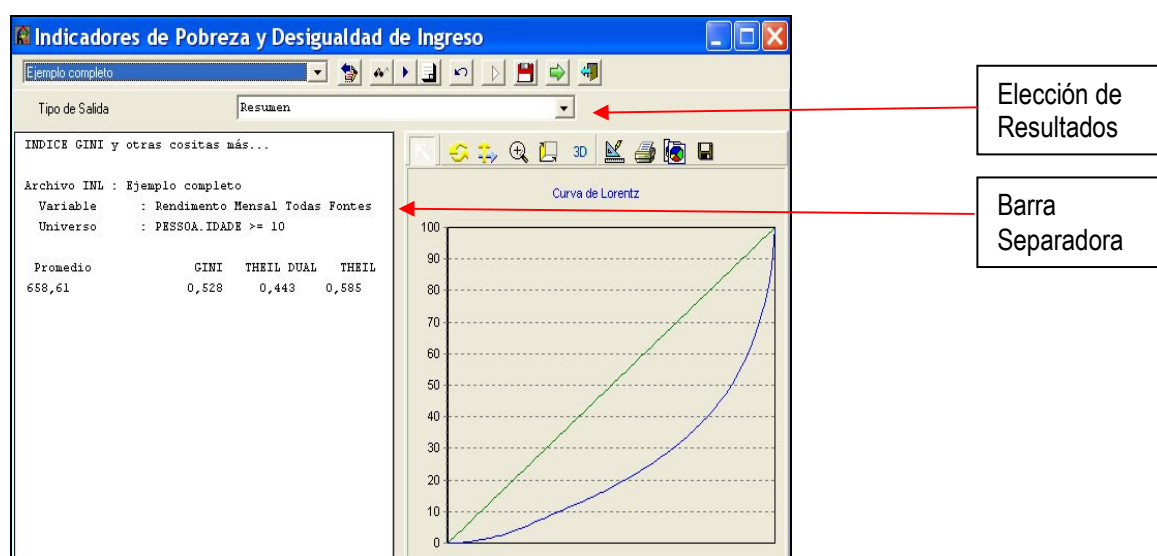


Figura 8 - Resultados Resumen

El panel de Resumen tiene dos partes, los resultados numéricos a la izquierda y la Curva de Lorenz a la derecha. El tamaño de estas partes puede ser modificado moviéndose la Barra Separadora para la izquierda o derecha. Si el área de salida y la selección geográfica tienen varios elementos (regiones, por ejemplo), sale una línea de resumen para cada región, y una Curva de Lorenz también para cada región.

La parte izquierda del panel con los resultados numéricos tiene un encabezado para identificar el archivo usado, la variable objetivo, los filtros aplicados a los datos, y una línea para cada área en la entidad de salida, con los valores del Promedio, el Índice de Gini, el Theil, y el Theil DUAL (ver VIII. Anexo 2 – Memoria de Cálculo).

La parte derecha del panel (la Curva de Lorenz) tiene algunos botones de control para la manipulación del gráfico. Los más importantes son los botones de Imprimir, Copiar, y Guardar el resultado. Es posible hacer un clic del Mouse sobre una de las curvas, y el programa muestra los resultados detallados para aquella curva.

Si una de las curvas mostradas aparece como que terminando antes del esperado, es porque esa área tiene menos casos que el número de puntos seleccionado para la ejecución. Eso no afecta mayormente los resultados numéricos, solamente la apariencia del gráfico.

VI.2 Resultados detallados

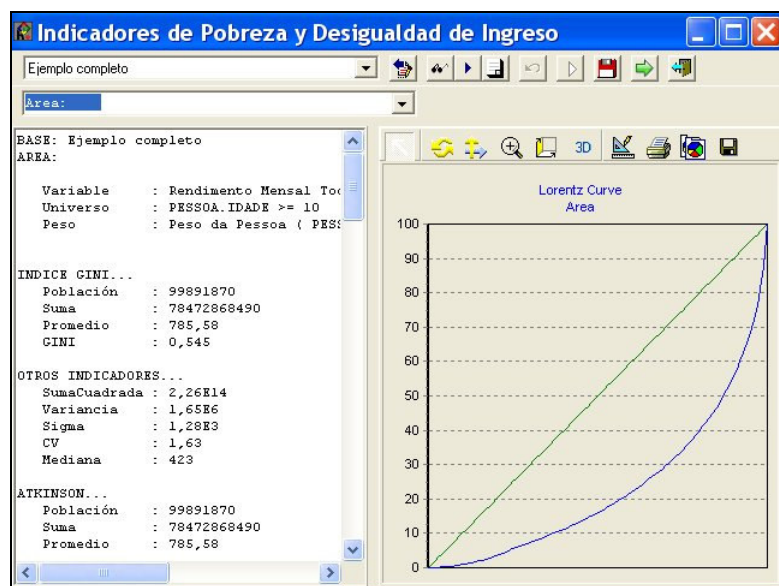


Figura 9 - Resultados Detallados

El programa muestra un panel detallado para cada área en la entidad de salida. Para la descripción de los resultados ver VIII. Anexo 2 – Memoria de Cálculo).

La parte inicial del panel trae la identificación del archivo INL con los parámetros, el área geográfica, la variable objetivo y los filtros usados. En seguida aparecen las informaciones del Índice de Gini, con algunos totales de apoyo, tales como la población, la suma de los valores, su promedio, y el Índice propiamente tal.

El programa muestra SIEMPRE los cálculos de los índices de Theil y Theil Dual, y los valores para Atkinson. Dependiendo de los parámetros en el archivo de entrada el programa muestra los detalles del Índice de Pobreza y la Descomposición del índice de Theil.

En el panel también se presenta, al final, la lista de los puntos calculados para la curva de Lorenz, con una serie de columnas de valores tales como el número de casos de, la suma de los valores, los porcentajes simples y acumulado, etc. A partir de estos valores es que se calculan los datos para la exportación al SIDRA.

VI.3 Programa REDATAM ejecutado

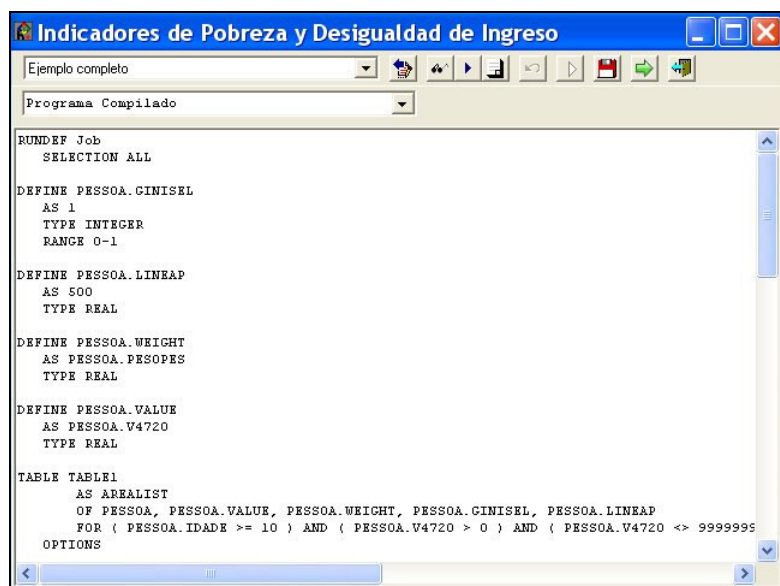


Figura 10 - Programa Ejecutado

Este panel trae el programa a ser enviado al Process de REDATAM para que sea ejecutado, y luego sus resultados son procesados en el Pós-Proceso del GINIPlus. Puede ser copiado y pegado en el Editor de Comandos del Process para que se entienda exactamente lo que hace el GINIPlus.

VII. Anexo 1 - Colores

La siguiente tabla trae la lista de nombres reservados de colores las cuales pueden ser usadas en el programa. Cualesquiera de los dos (Nombre o Alias) pueden ser usados, escritos siempre en MAYÚSCULAS. Si el nombre no es reconocido por el programa será remplazado por el color BLANCO. Los colores también pueden ser informados por un número en formato RGB.

| Nombre | Alias |
|--------------|----------|
| | |
| AQUA | AGUA |
| WHITE | BLANCO |
| DARKGRAY | |
| GRAY | |
| FUCHSIA | FUCCIA |
| BLACK | NEGRO |
| BLUE | AZUL |
| GREEN | VERDE |
| RED | ROJO |
| LIME | LIMA |
| MAROON | |
| NAVY | |
| YELLOW | AMARILLO |
| TEAL | |
| SILVER | PLATA |
| OLIVE | OLIVA |
| PURPLE | PURPURA |
| LENIN | |
| DARKLENIN | |
| CHILEINE | |
| DARKCHILEINE | |
| SILVERLITE | |

Ejemplos:

```
COLOR=LIMA
COLORMIN=128.128.0
COLORMAX=DARKLENIN
```

VIII. Anexo 2 – Memoria de Cálculo

Para más detalles, ver “Estatísticas para Economistas”, de Rodolfo Hoffmann. Editora Thomson Learning Edições Ltda.

VIII.1 Curva de Lorenz y el Índice de Gini

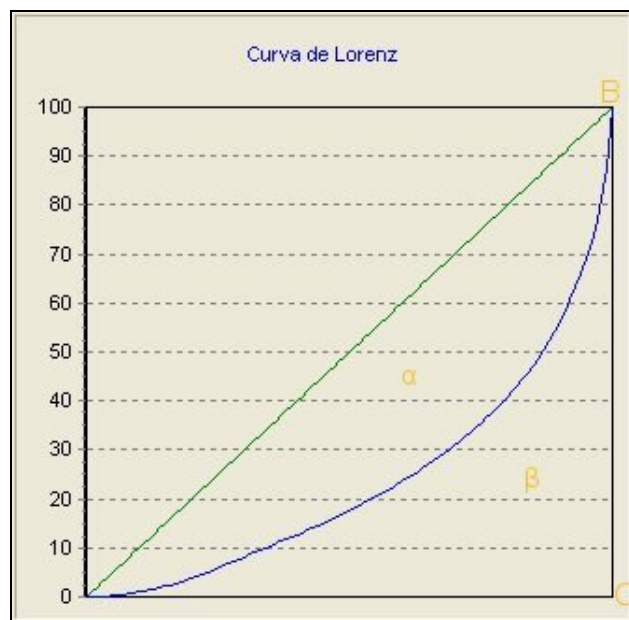


Figura 11 - Curva de Lorenz

Para la construcción de la Curva de Lorenz (en azul) se organizan los casos en orden creciente del valor de la variable objetivo (por ejemplo, el valor del ingreso). Se acumulan en el eje del X los porcentajes de casos, y se acumulan en el eje del Y los porcentajes de la variable objetivo.

El área “ α ”, denominada de “área de concentración”, está comprendida entre la curva de Lorenz y el segmento OB.

Una distribución de ingreso con una perfecta igualdad estaría representada por una “Curva” de Lorenz reducida a un segmento de recta OB, también llamada de “línea de la perfecta igualdad”. En este caso el área de concentración α sería 0.

Por otro lado, una distribución de ingreso con un máximo de desigualdad sería el caso de una población de n individuos en donde $n-1$ no reciben ingreso y un individuo recibe el ingreso total. En este caso, la proporción acumulada de ingreso es cero hasta el punto $(n-1)/n$, pasando a 1 cuando se incluye el último individuo (el de ingreso total). Para una población bastante grande, la “curva” de Lorenz se confunde con la línea poligonal OCB, y el área de concentración sería el área del triángulo OCB, o sea, 0,5.

Por definición, el Índice de Gini (G) es la relación entre el área de concentración α y el área del triángulo OCB, o sea

$$G = \alpha / 0,5 = 2 \alpha$$

Una vez que $0 \leq \alpha \leq 0,5$, entonces $0 \leq G \leq 1$. Es decir, el Índice de Gini varía de 0 (igualdad total) a 1 (máxima desigualdad)

VIII.2 Índice de Theil

En el programa, por convención, se llama la Redundancia de Índice de Theil, y se llama el Índice de Theil de Theil DUAL. La Redundancia tiene la fórmula

$$R = \sum y * \log (y / n)$$

En el caso de la perfecta igualdad $R = 0$, y si hay una perfecta desigualdad, $R = \log (n)$

El Theil Dual es calculado por

$$T = 1 - \exp (-R)$$

La descomposición en grupos (estratos) es calculada para evaluar la influencia de una variable (sexo, escolaridad, etc.) en la desigualdad. Ella tiene dos partes, la desigualdad entre los grupos (R1), y la desigualdad dentro de cada grupo (R2).

$$R = R1 + R2$$

$$R1 = \sum y_g * \log (y_g / n_g)$$

$$R2 = \sum y_g * \{ \sum Y_i / Y_g \log [(Y_i / Y_g) / (1 / N_g)] \}$$

| | |
|-------|--------------------------------------|
| y_g | Porcentaje de ingreso en el grupo g |
| n_g | Porcentaje de personas en el grupo g |
| Y_i | Ingreso de la persona en el grupo g |
| Y_g | Ingreso total del grupo g |
| N_g | Número de personas en el grupo g |

VIII.3 Índice de Atkinson

Akinson (Atk) es uno de los índices basados en la función de bienestar social. Las fórmulas de la familia de índices normativos de Atkinson están bien descritas en “DESIGUALDAD, DIVERSIDAD Y CONVERGENCIA: (ALGUNOS) INSTRUMENTOS DE MEDIDA”, de Francisco J. Goerlich.

El programa calcula el Atk para una serie de valores para el parámetro ϵ (0, 0.5, 1, 1.5, 2 y 3). El parámetro ϵ se interpreta como el grado de aversión relativa a la desigualdad por parte de la sociedad y se supone constante. Para $\epsilon = 0$ no existe ninguna aversión social a la desigualdad. A medida que aumenta ϵ aumenta el grado de aversión social a la desigualdad.

VIII.4 Línea de Pobreza

Sea L el valor de la Línea de Pobreza informada como parámetro de entrada del programa, y que la variable objetivo sea el ingreso de las personas. A partir de este valor L se calculan el número de pobres (casos en los cuales el valor de la variable objetivo está por debajo de ese valor L), el ingreso promedio, y los indicadores a seguir.

| Valores básicos | |
|-----------------|---|
| L | Línea de pobreza informada por el usuario |
| x | Variable de ingreso |
| n | Población total en el archivo |
| k | Población pobre (personas con ingreso $< L$) |
| m | Ingreso promedio de los pobres ($m = \sum x / k$) |
| mu | Ingreso promedio total ($\mu = \sum x / n$) |
| totrich | Sumatoria del ingreso total de los ricos |
| totrich1 | Sumatoria del ingreso total de los ricos descontado el valor de la línea de pobreza |

| Indicadores calculados | |
|------------------------|---|
| G | Índice de Gini de los pobres (personas con $x < L$) |
| H | Proporción de pobres en la población ($H = k / n$) |
| I | Insuficiencia de ingreso ($I = (L - m) / L$) |
| HI | Combinación de H e I ($HI = H * I$). En cuanto H es insensible a la intensidad de la pobreza, I es insensible al número de pobres, por lo que se combinan las dos medidas con el indicador HI . |
| P0 | Igual que H |
| P1 | Igual que HI |
| P2 | Combinación de H e I pero tomándose I^2 |
| P3 | Combinación de H e I pero tomándose I^3 |
| P | Índice de Sen ($P = H [I + (1 - I) G]$) |
| P1_0 | Proporción del valor total en cuanto debería crecer el ingreso de los pobres hasta la línea de pobreza en relación al ingreso total ($k * (L - m) / (n * \mu)$) |
| P2_0 | Mismo numerador de P1_0, pero en relación al ingreso total de la población "rica" ($k * (L - m) / (\text{totrich})$) |
| P3_0 | Mismo numerador de P1_0, pero en relación al ingreso total de la población rica descontado el valor de la Línea de Pobreza ($k * (L - m) / (\text{totrich1})$) |