

# REVISTA ESTADISTICA

Vol. 65 - NRO. 184 – junio 2013

Número Especial: 70 años de la revista Estadística

Special Issue: 70 years of the Journal Estadística

## CONTENIDO – CONTENTS

### **THE BETA-G FAMILY OF DISTRIBUTIONS**

GAUSS M. CORDEIRO

*Universidade Federal de Pernambuco, Brazil*  
[gausscordeiro@uol.com.br](mailto:gausscordeiro@uol.com.br)

ARTUR J. LEMONTE

*Universidade Federal de Pernambuco, Brazil*  
[arturlemonte@gmail.com](mailto:arturlemonte@gmail.com)

EDWIN M.M. ORTEGA

*Universidade de São Paulo, Brazil*  
[edwin@usp.br](mailto:edwin@usp.br)

### **ABSTRACT**

The beta-G family of distributions is the subject of this paper. This family includes all the exponentiated distributions. Further, it allows for greater flexibility of its tails and can be widely applied in many areas such as engineering, biology, medical, and so forth. In the last years, several

new beta-G models have been proposed, mostly by statisticians in Brazil. So, we feel that it is timely for a review of the beta-G family of distributions. The aim of this paper is just that.

**Keywords**

Beta distribution; Exponentiated distribution; Extended distribution; Maximum likelihood estimation.

**RESUMEN**

Este artículo trata sobre la familia Beta-G de distribuciones. Esta familia incluye a todas las distribuciones exponenciales. Más aún, permite mayor flexibilidad en sus colas y se puede aplicar en muchas áreas tales como ingeniería, biología, medicina y otras. En los últimos años, se propusieron nuevos modelos Beta-G, en su mayoría fueron realizados por estadísticos de Brasil. Entonces, creemos que es oportuno realizar una revisión de la familia Beta-G de distribuciones. Ese es básicamente el objetivo de este artículo.

**Palabras clave**

Distribución Beta; Distribución a exponentes; Distribución extendida; Estimación de máxima verosimilitud.

**COORDINATION OF STATISTICS IN A DECENTRALIZED  
STATISTICAL SYSTEM: INSTRUMENTS AND POLICIES**

JACOB RYTEN

*International Consultant, Ex-President of IASI*

[jacob.ryten@gmail.com](mailto:jacob.ryten@gmail.com)

**ABSTRACT**

This paper complements United Nations' *Handbook of Statistical Organization, Third Edition. The Operation and Organization of a Statistical Agency* (2003). The latter analysed the mechanisms underlying centralized official statistical agencies. The current paper shifts the argument to decentralized systems where there is no machinery to ensure coordination among the component partners. A strong criticism is made of the frequently encountered attempt to ensure coordination by means of a paper "national statistical system" (NSS). There is no single recorded success story or even one of measurable progress secured because of a NSS. Rather than relying on a meaningless agreement, the government statistical office should analyse those of its systems that can be used to advantage by other government agencies ready to enter into a partnership and use without violating any of the provisions of the Statistics Act such dedicated instruments as the official lists of

businesses, Census derived urban cartography, and above all expert knowledge of sampling techniques. Only by building slowly but progressively a reputation of technical competence and after creating an impressive record of thoughtful delivery of results, can the statistical office move from a marginal rather than central location. Moreover such a move is essential to allow it to take on a more strategic role and eventually become the authority to certify as official the results of statistical inquiries and administrative records conducted and held by its partners in government.

***Key words***

Data coherence; statistical systems; centralized systems.

**RESUMEN**

Este artículo sirve de complemento al *Manual de organización estadística, tercera edición. El funcionamiento y la organización de una oficina estadística* de Naciones Unidas (2003). Ese manual fue dedicado al análisis de los mecanismos con que operan las oficinas de estadística centralizadas. Pero este artículo cambia de énfasis para examinar los sistemas descentralizados en que típicamente no existen mecanismos para asegurar la coordinación entre los distintos miembros. Se critica el intento común a muchos países de crear un Sistema Estadístico Nacional (SEN) basado exclusivamente en un decreto o resolución sin el apoyo de mecanismos de certificación y de coacción concretos. Sin embargo no existe un solo ejemplo documentado de éxito logrado a partir de tales sistemas. El artículo sugiere que la oficina de estadística suministre paulatinamente a los demás miembros del sistema, pero en forma progresiva, los instrumentos que le dan una ventaja comparativa en la solución de problemas técnicos sin comprometer sus normas legales tales como el secreto estadístico. En particular, la utilización de directorios de empresas, la cartografía urbana censal y las técnicas de muestreo pueden permitir que la oficina de estadística ocupe un lugar estratégico en las deliberaciones del gobierno que demandan información cuantitativa. Una evolución continua y progresiva llevará la oficina de estadística a ocupar un lugar más central y en un futuro no muy lejano, a ser la agencia certificadora de encuestas estadísticas y de registros administrativos actualmente en manos de otros ministerios.

***Palabras clave***

Coherencia de datos; sistemas estadísticos; sistemas centralizados.

**THANKS TO 70 YEARS OF INTER AMERICAN  
STATISTICAL COOPERATION, THE WORLD'S LARGEST  
INTEGRATED CENSUS MICRODATA DISSEMINATION  
SITE [WWW.IPUMS.ORG/INTERNATIONAL](http://WWW.IPUMS.ORG/INTERNATIONAL)**

ROBERT MCCA  
*University of Minnesota Population Center, USA*  
[rmccaa@umn.edu](mailto:rmccaa@umn.edu)

## ABSTRACT

Seventy years of Inter American Statistical cooperation, symbolized by the 70<sup>th</sup> anniversary of *Estadística*, made possible the construction of IPUMS-International, the world's largest integrated census microdata dissemination site, [www.ipums.org/international](http://www.ipums.org/international). Currently, the site offers access to 238 samples totaling over 540 million person records representing 74 countries. The Americas, which account for only about one-seventh of the world's population, amount to over one-third (36%) of the person records in the IPUMS-International database. Likewise, 35% of the citations in the IPUMS-International bibliography are for studies focused on Latin America, with about half of these analyzing a single Latin American country. This article discusses salient features of the IPUMS integration methods and system. National Statistical Institutes that have not yet entrusted 2010 census microdata to the initiative are invited to do so. Researchers and teachers are invited to use the data freely in analysis and teaching.

### *Key words:*

Census, sample, public statistics, microdata, metadata, dissemination, IPUMS

## RESUMEN

Setenta años de cooperación estadística interamericana, simbolizada por el 70º aniversario de la revista *Estadística*, han hecho posible la construcción de IPUMS internacional, la base en línea de microdatos censales armonizados más grande del mundo, [www.ipums.org/international](http://www.ipums.org/international). Actualmente, IPUMS proporciona acceso a muestras con más de 540 millones de registros individuales de 74 países. Las Américas, que albergan una séptima parte de la población mundial, representan más de un tercio (36%) de todos los registros individuales en la base de datos IPUMS-internacional. Asimismo, el 35% de todas las referencias en la bibliografía de IPUMS son de estudios realizados sobre América Latina, la mitad de éstas basadas en un sólo país de la región. Este artículo presenta las principales características del sistema de integración y difusión de datos de IPUMS. Los Institutos Nacionales de Estadística que todavía no han entregado la muestra de microdatos de la ronda de 2010 son invitados a hacerlo. Los investigadores y profesores son invitados a utilizar los datos de forma gratuita para sus actividades de investigación y docencia.

### *Palabras clave:*

Censo, muestra, estadística pública, microdatos, metadatos, difusión, IPUMS

## **SAMPLE DESIGN FOR THE NUTRITION, PHYSICAL ACTIVITY AND HEALTH SURVEY (PNAFS), NITERÓI, RIO DE JANEIRO, BRAZIL**

MAURICIO TEIXEIRA LEITE DE VASCONCELLOS  
*Escola Nacional de Ciências Estatísticas*  
*Fundação Instituto Brasileiro de Geografia e Estatística*  
[mautlv@gmail.com](mailto:mautlv@gmail.com)

PEDRO LUIS DO NASCIMENTO SILVA  
*Escola Nacional de Ciências Estatísticas*  
*Fundação Instituto Brasileiro de Geografia e Estatística*  
[pedronsilva@gmail.com](mailto:pedronsilva@gmail.com)

LUIZ ANTONIO DOS ANJOS  
*Laboratório de Avaliação Nutricional e Funcional*  
*Departamento de Nutrição Social, Universidade Federal Fluminense*  
[lanjos@gmail.com](mailto:lanjos@gmail.com)

### ABSTRACT

This paper reports the methods used to select and weight the sample and subsamples used in the Nutrition, Physical Activity and Health Survey, a household survey conducted in 2003 by the Nutritional and Functional Assessment Laboratory of the Fluminense Federal University. The sample was selected in three stages. In the first, census enumeration areas were selected with probability proportional to the number of private households. In the second, households were selected using an inverse sampling procedure. In the third, one healthy adult was selected to be monitored in real life conditions and subsamples of adults were invited to come to the laboratory for a series of physical and physiological measurements. Initially, the sampling weights were calculated as the reciprocal of the product of the inclusion probabilities in each sampling stage and then were calibrated to maintain coherence to population totals obtained from other sources.

#### *Key words*

Household sample; inverse sampling; integrated household weighting method

### RESUMEN

Este artículo presenta los métodos utilizados para seleccionar y ponderar la muestra y submuestras utilizadas en la Encuesta de Nutrición, Actividad Física y de Salud, una encuesta de hogares realizada en 2003 por el Laboratorio de Evaluación Nutricional y Funcional de la Universidad Federal Fluminense. La muestra fue seleccionada en tres etapas. En la primera, las zonas de empadronamiento censal se seleccionaron con probabilidad proporcional al número de hogares. En la segunda, los hogares fueron seleccionados mediante un procedimiento de muestreo inverso. En la tercera, un adulto saludable fue seleccionado para ser monitoreado en condiciones reales; y submuestras de adultos fueron invitados para asistir al laboratorio para realizar una serie de mediciones físicas y fisiológicas. Inicialmente, las ponderaciones muestrales se calcularon como el recíproco del producto de las probabilidades de inclusión en cada etapa de muestreo y luego se calibraron para mantener la coherencia con los totales de población obtenidos a partir de otras fuentes.

#### *Palabras Clave*

Muestra de hogares; muestreo inverso; método integrado de ponderación de Hogares

# TEST OF OUTLIERS AND INFLUENTIAL OBSERVATIONS IN GARCH MODELS: A REVIEW

LUIZ K. HOTTA

*Department of Statistics, University of Campinas, Brazil*  
[hotta@ime.unicamp.br](mailto:hotta@ime.unicamp.br)

MAURICIO ZEVALLOS

*Department of Statistics, University of Campinas, Brazil*  
[amadeus@ime.unicamp.br](mailto:amadeus@ime.unicamp.br)

## ABSTRACT

When modeling time series, the detection of atypical observations that cannot be accommodated by a model is a crucial part. These observations might produce strong effects on the statistical analysis. For this reason, the study of detection methods is an important research topic. The objective of this paper is to present a review of proposed methods to detect outliers and influential points in generalized autoregressive conditional heteroscedastic (GARCH) models, which are routinely used for estimating the volatility of financial time series.

### *Keywords*

Level outlier; volatility outlier; additive outlier; innovation outlier; detection of outliers.

## RESUMEN

En el proceso de construcción de modelo para series de tiempo, es crucial la detección de observaciones atípicas, las cuales no pueden ser explicadas por el modelo. Estas observaciones podrían afectar severamente el análisis estadístico. Por este motivo, el estudio de métodos de detección constituye un tema importante de investigación. El objetivo de este artículo es presentar una revisión de métodos que han sido propuestos para detectar outliers y puntos influyentes en modelos autoregresivos de varianza condicional heterocedástica (GARCH), modelos que son rutinariamente utilizados para estimar la volatilidad de series financieras.

### *Palabras Clave*

Outlier de nivel; outlier de volatilidad; outlier aditivo; outlier de innovación; detección de outliers.