

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE MATEMÁTICA

MAESTRÍA EN MODELAJE MATEMÁTICO
PRIMERA VERSIÓN



UN INDICADOR DINÁMICO BASADO EN
COMPONENTES PRINCIPALES COMO
ALTERNATIVA AL ÍNDICE DE NECESIDADES
BÁSICAS INSATISFECHAS EN LOS CENSOS.
CASO DE ESTUDIO: BOLIVIA

TRABAJO DE GRADO PARA LA OBTENCIÓN DEL TÍTULO DE
MAGISTER EN CIENCIAS EN MODELAJE MATEMÁTICO

Autor: Lic. Alvaro Limber Chirino Gutierrez

Tutor: M. Sc. Charlie Lozano Correa

LA PAZ - BOLIVIA
2019

Universidad Mayor de San Andrés
Facultad de Ciencias Puras y Naturales
Carrera de Matemática
Maestría en Modelaje Matemático Primera version

Tesis de Maestría:

Un indicador dinámico basado en componentes principales como alternativa al Índice de Necesidades Básicas Insatisfechas en los Censos. Caso de estudio: Bolivia

Para optar el grado Académico de Magister Scientiarum en Modelaje Matemático del
Postulante:

Alvaro Limber Chirino Gutierrez

Nota Numeral:

Nota Literal:

Significado de la calificación:

Director Académico:

Co - Director:

Tutor:

Tribunal:

Tribunal:

Resumen

El índice de necesidades básicas insatisfechas (NBI) es uno de los indicadores más comunes en América Latina al momento de aproximar el bienestar de una población una vez realizado un censo de población y vivienda. El NBI está definido en base a variables propias de un censo (limitación temática) y pretende aproximarse al bienestar de los hogares definiendo en su fórmula dimensiones de interés y ponderaciones subjetivas de esas áreas. Tanto las variables que se emplean como las ponderaciones son invariantes en el tiempo.

El problema central es que no existe un indicador de bienestar asociado a los censos que esté libre de subjetividad respecto a las variables que lo definen y las ponderaciones que se dan. La limitación principal de esta concepción es que se define de manera estática en el tiempo al concepto de bienestar. Se propone construir un indicador que inicialmente tome en cuenta un set de variables disponibles en el censo y aplique sobre ellas una transformación mediante el uso de componentes principales.

El trabajo plantea un indicador de riqueza basado en el uso de componentes principales para un set completo de variables binarias correspondientes a; características de la vivienda y el hogar, servicios y activos del hogar e información geográfica. Donde se postula que el primer componente asociado a un set de variables optimizadas es un indicador que permite monitorear el bienestar como una alternativa al NBI, teniendo como ventajas la identificación de las variables más relevantes e identificando ponderaciones para ellas.

El trabajo se desarrolla dentro del área de Estadística Multivariante, se emplean los datos de los tres últimos censos de Población y Vivienda de Bolivia; el CNPV-1992, el CNPV-2001 y el CNPV-2012, se incluyen los tres censos con la finalidad de explorar los cambios del indicador de riqueza, basados en un periodo base de definición del indicador. Para el Censo 2012 se construye el indicador a nivel departamental y se contrasta entre departamentos.

Para el censo de 1992 se inició con 63 variables y mediante los pasos definidos se identificó a 17 variables relevantes para la definición del indicador de riqueza, para el censo de 2001, se partió con 77 variables y se conservaron 17 variables, en el caso del censo de 2012, se inició con 84 variables y se concluyó con 22 variables seleccionadas. Se identifica la funcionalidad del indicador y el aporte de cada variable seleccionada. Tomando en cuenta el censo de 1992 como base, se ve el cambio en el tiempo con los otros censos y se puede identificar los departamentos que presentan mayores cambios, de igual forma para el censo de 2001. Tanto Santa Cruz y Tarija son los departamentos con mejores niveles en riqueza y mejor avance en el indicador de riqueza.

El trabajo presenta una propuesta metodológica basada en criterios objetivos y que permite identificar las variables relevantes para aproximar al bienestar, otorgando al mismo tiempo una base para que el monitoreo del bienestar en base a un periodo definido, donde la única condición es tomar en cuenta las variables definidas por el indicador de riqueza. Esta metodología planteada puede adaptarse fácilmente a otros ámbitos, este

trabajo estuvo concentrado en variables de la vivienda y el hogar vinculadas a un proxis de riqueza.

Abstract

The index of unsatisfied basic needs (UBN) is one of the most common indicators in Latin America when it comes to approximate the wellbeing of a population once a population and housing census has been taken. The NBI is defined based on variables of a census (thematic limitation) and aims to approximate the wellbeing of households by defining in their formula dimensions of interest and subjective weights of those areas. Both the variables that are used and the weights are invariant over time.

The central problem is that there is no indicator of welfare associated with censuses that is free of subjectivity with respect to the variables that define it and the weightings that are given. The main limitation of this conception is that the concept of wellbeing is defined in a static way over time. It is proposed to construct an indicator that initially takes into account all the variables available in a census and apply a transformation to them through the use of principal components.

The paper presents a wealth indicator based on the principal components method for a complete set of binary variables corresponding to; housing characteristics and household services and household assets and geographic information. Where it is postulated that the first component associated with a set of optimized variables is an indicator that allows monitoring well-being as an alternative to the UBN, having as advantages the identification of the most relevant variables and identifying weights for them.

The work takes place in the framework of Multivariate Statistics, the data correspond to the last three of National Population and Housing Census of Bolivia are used; the *NPHC* – 1992, the *NPHC* – 2001 and the *NPHC* – 2012, the three census are included in order to explore the changes of the wealth indicator, based on a base period of definition of the indicator. For the 2012 Census the indicator is built at the departmental level and contrasted between departments.

For the 1992 census, 63 variables were started and through the defined steps, 17 relevant variables were identified for the definition of the wealth indicator, for the 2001 census, it was split with 77 variables and 17 variables were conserved, in the case of 2012 census, started with 84 variables and concluded with 22 selected variables. The functionality of the indicator and the contribution of each selected variable are identified. Taking into account the 1992 census as a basis, the change in time is seen with the other censuses and the departments that present the greatest change can be identified, in the same way for the 2001 census. Both Santa Cruz and Tarija are the departments with better levels of wealth and better progress in the wealth indicator.

The paper presents a methodological proposal based on objective criteria that allows identifying the relevant variables to approximate well-being, while providing a basis for monitoring welfare based on a defined period, where the only condition is to take into account the variables defined by the wealth indicator. This proposed methodology can be easily adapted to other areas, this work was concentrated in housing and household variables linked to a wealth proxies.

Índice general

1	Generalidades	7
1.1	Introducción	7
1.2	Antecedentes	7
1.3	Justificación	9
1.4	Planteamiento del problema	9
1.5	Objetivos	9
1.5.1	Objetivo General	9
1.5.2	Objetivos específicos	10
1.6	Operacionalización de los objetivos de estudio	10
1.7	Alcances	11
2	Marco Teórico	12
2.1	Agenda Patriótica, Plan Nacional de Desarrollo y Objetivos de Desarrollo Sostenible	12
2.2	El Índice de Necesidades Básicas Insatisfechas (NBI) y sus limitaciones	14
2.2.1	El método de NBI	14
2.2.2	Críticas y limitaciones al método del NBI	16
2.3	Alternativas al NBI	16
2.4	Métodos Multivariantes	17
2.4.1	La organización de los datos	18
2.4.2	Álgebra de Matrices y Vectores aleatorios	19
2.5	Método de componentes principales	22
2.5.1	Eligiendo un sub conjunto de componentes principales	26
2.5.2	Eligiendo un sub conjunto variables	28
2.5.3	Un ejemplo de CPs sobre medidas de satisfacción	29
2.6	Componentes Principales para datos Binarios	31
2.6.1	Un ejemplo con la Encuesta a Hogares 2015, relación entre la pobreza y el indicador basado en CPs	32
3	Metodología de la investigación	36
3.1	Diseño de la investigación	36
3.2	Pasos para la definición del “indicador de Bienestar”	36
3.3	CPs como un indicador de monitoreo	37
3.4	Datos y dominios	38
3.5	Variables	39

4	Resultados, conclusiones y recomendaciones	42
4.1	Resultados	42
4.1.1	CNPV 1992 vs CNPV 2001 y CNPV 2012	43
4.1.2	CNPV 2001 vs CNPV 2012	49
4.1.3	CNPV 2012	56
4.1.4	Para el CNPV 2012, a nivel de los 9 departamentos	63
4.2	Conclusiones, recomendaciones y limitaciones	69
5	Propuesta de Mejoramiento	70
A	Anexos	71
A.1	Siglas	71
A.2	Eigenvalores y eigenvectores finales para el censo 2012 por departamentos	72
	Bibliografía	77

Índice de cuadros

2.1	Resultados para la meta 1 del pilar: Erradicar la pobreza extrema	14
2.2	Calificación para el material predominante en el techo	16
2.3	Matriz de correlación (ρ) para las 5 variables de satisfacción, $N = 9147$.	29
2.4	Eigenvalores y eigenvectores de la matriz de correlación de las variables de satisfacción y el <i>PVT</i>	30
2.5	Matriz de cargas factoriales (correlaciones <i>CPs</i> y variables) para las variables de satisfacción	30
2.6	Matriz de varianzas y covarianzas para las 10 variables de equipamiento .	33
2.7	Eigenvalores y eigenvectores de la matriz de covarianza de las variables de equipamiento y el <i>PVT</i>	33
2.8	Correlaciones entre los 2 primeros componentes y las variables de equipamiento	34
2.9	Modelo 1	34
2.10	Modelo 2	35
2.11	Modelo 3	35
3.1	Variables a nivel de vivienda y hogar en los Censos de Población y Vivienda 1992, 2001 y 2012	41
4.1	Variables y ponderadores para el indicador de riqueza según el censo base	42
4.2	Eigenvalores y % acumulado para el indicador de riqueza, 1992	47
4.3	Eigenvectores para el indicador de riqueza, 1992	47
4.4	Eigenvalores y % acumulado para el indicador de riqueza, 2001	54
4.5	Eigenvectores para el indicador de riqueza, 2001	54
4.6	Eigenvalores y % acumulado para el indicador de riqueza, 2012	61
4.7	Eigenvectores para el indicador de riqueza, 2012	61
A.1	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Chuquisaca	72
A.2	Eigenvalores y % acumulado para el indicador de riqueza, 2012. La Paz .	72
A.3	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Cochabamba	72
A.4	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Oruro . .	73
A.5	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Potosí . .	73
A.6	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Tarija . .	73
A.7	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Santa Cruz	73
A.8	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Beni . . .	73
A.9	Eigenvalores y % acumulado para el indicador de riqueza, 2012. Pando . .	73
A.10	Eigenvectores para el indicador de riqueza, 2012. Chuquisaca	73
A.11	Eigenvectores para el indicador de riqueza, 2012. La Paz	74
A.12	Eigenvectores para el indicador de riqueza, 2012. Cochabamba	74
A.13	Eigenvectores para el indicador de riqueza, 2012. Oruro	74

A.14	Eigenvectores para el indicador de riqueza, 2012. Potosi	75
A.15	Eigenvectores para el indicador de riqueza, 2012. Tarija	75
A.16	Eigenvectores para el indicador de riqueza, 2012. Santa Cruz	75
A.17	Eigenvectores para el indicador de riqueza, 2012. Beni	76
A.18	Eigenvectores para el indicador de riqueza, 2012. Pando	76

Índice de figuras

2.1	Necesidades Básicas, dimensiones y variables censales	15
2.2	Ejemplo de un scree plot para datos simulados	27
4.1	Correlaciones entre las primeras 63 variables del censo de 1992	43
4.2	Scree Plot del componente principal inicial para el censo de 1992	43
4.3	Porcentaje acumulada de la varianza para el componente inicial para el censo de 1992	44
4.4	Correlaciones entre las 17 variables resultantes de los criterios, para el censo de 1992	45
4.5	Scree Plot del componente principal inicial para el censo de 1992	45
4.6	Porcentaje acumulada de la varianza para el componente inicial para el censo de 1992	46
4.7	Eigenvalores y dirección para el primer componente para el censo de 1992	46
4.8	Densidad e histograma a partir del indicador para 1992	47
4.9	Diagramas de caja por Censo del indicador de riqueza	48
4.10	Diagrama de caja por Censo y área del indicador de riqueza	48
4.11	Diagrama de caja por Censo y departamento del indicador de riqueza . .	48
4.12	Evolución del indicador de riqueza a nivel municipal en base al censo de 1992	49
4.13	Correlaciones entre las primeras 77 variables del censo de 2001	50
4.14	Scree Plot del componente principal inicial para el censo de 2001	50
4.15	Porcentaje acumulada de la varianza para el componente inicial para el censo de 2001	51
4.16	Correlaciones entre las 18 variables resultantes de los criterios, para el censo de 2001	52
4.17	Scree Plot del componente principal inicial para el censo de 2001	52
4.18	Porcentaje acumulada de la varianza para el componente inicial para el censo de 2001	53
4.19	Eigenvalores y dirección para el primer componente para el censo de 2001	53
4.20	Densidad e histograma a partir del indicador para 2001	54
4.21	Diagrama de caja por Censo del indicador de riqueza	55
4.22	Diagrama de caja por Censo y área del indicador de riqueza	55
4.23	Diagrama de caja por Censo y departamento del indicador de riqueza . .	55
4.24	Evolución del indicador de riqueza a nivel municipal en base al censo de 2001	56
4.25	Correlaciones entre las primeras 84 variables del censo de 2012	57
4.26	Scree Plot del componente principal inicial para el censo de 2012	57
4.27	Porcentaje acumulada de la varianza para el componente inicial para el censo de 2012	58

4.28	Correlaciones entre las 22 variables resultantes de los criterios, para el censo de 2012	59
4.29	Scree Plot del componente principal inicial para el censo de 2012	59
4.30	Porcentaje acumulada de la varianza para el componente inicial para el censo de 2012	60
4.31	Eigenvalores y dirección para el primer componente para el censo de 1992	60
4.32	Densidad e histograma a partir del indicador para 2012	61
4.33	Diagrama de caja del indicador de riqueza, Censo 2012	62
4.34	Diagrama de caja por área del indicador de riqueza, Censo 2012	62
4.35	Diagrama de caja por departamento del indicador de riqueza, Censo 2012	62
4.36	Indicador de riqueza a nivel municipal en base al censo de 2012	63
4.37	Correlaciones entre las variables del indicador de riqueza, por departamen- to. Censo 2012	64
4.38	Scree Plot del componente principal para el indicador de riqueza, por de- partamento. Censo 2012	65
4.39	Porcentaje acumulado de la varianza para el componente principal para el indicador de riqueza, por departamento. Censo 2012	66
4.40	Eigenvalores y dirección para el primer componente del indicador de rique- za, por departamento. Censo 2012	67
4.41	Diagramas de caja según el indicador de riqueza, base a nivel departamental	68

Capítulo 1

Generalidades

1.1 Introducción

El trabajo se organiza en capítulos, en **el capítulo 1** se presentan las generalidades del trabajo, en este se desarrolla los antecedentes, la justificación, el planteamiento del problema, los objetivos, la operacionalización de los objetivos y los alcances del trabajo. En **el capítulo 2** se incluye el marco teórico del trabajo, éste aborda; la agenda patriótica, el índice de necesidades básicas insatisfechas (NBI), las alternativas al NBI, los métodos multivariantes, describiendo el método de componentes principales para datos binarios junto con las consideraciones en su tratamiento. En **el capítulo 3** se desarrolla la metodología de la investigación, definiendo el diseño de la investigación, se listan los pasos para la definición del indicador de riqueza y se describen las fuentes de datos a usar junto con las variables vinculadas. En **el capítulo 4** se presentan los resultados, las conclusiones y recomendaciones del estudio. Finalmente, en **el capítulo 5** se presenta la propuesta de mejoramiento a partir de los hallazgos del trabajo.

El presente trabajo de grado consideró en su estructura y elaboración la **Guía Metodológica para la elaboración del trabajo de grado. Programas de maestría PGAMAT-IIMAT** y el **Reglamento de tesis de grado PGAMAT-IIMAT**. Siguiendo el artículo 8, del reglamento:

ARTÍCULO 8. La tesis de grado es la culminación de un trabajo de investigación aplicado, original cuyo valor científico, artístico o literario deberá ser congruente con el grado académico que se pretende alcanzar.

Este documento desarrolla un trabajo de investigación aplicado, el mismo fue desarrollado considerando el contenido de la Maestría y en particular los temas vistos en las asignaturas de: Matrices, Análisis de Datos, Minería de datos y Análisis de Métodos Multivariantes.

1.2 Antecedentes

El índice de necesidades básicas insatisfechas (NBI) es uno de los indicadores más comunes al momento de aproximar el bienestar de una población una vez realizado un censo de población y vivienda. El NBI está definido en base a variables propias de un censo (limitación temática) y pretende aproximarse al bienestar de los hogares definiendo en

su fórmula, dimensiones de interés y ponderaciones subjetivas de esas áreas. Tanto las dimensiones que se emplean como las ponderaciones son invariantes en el tiempo¹.

Tradicionalmente el NBI considera las dimensiones de²:

1. Acceso a la vivienda
 - Calidad de la vivienda
 - Hacinamiento
2. Acceso a los servicios sanitarios
 - Disponibilidad de agua potable
 - Tipo de sistema de eliminación de excretas
3. Acceso a la educación
 - Asistencia de los niños en edad escolar a un establecimiento educativo
4. Capacidad Económica
 - Probabilidad de insuficiencia de ingresos en el hogar

Sobre las cuales se aplican normas y ponderaciones para la construcción del indicador global. Sin embargo, las poblaciones son dinámicas en el tiempo, por ejemplo, si en cierto año todos los hogares tienen el mismo tipo de piso o disponibilidad de servicio sanitario, esa variable no está aportando información al indicador.

La importancia del NBI radica que este indicador contribuye al monitoreo de la Agenda Patriótica, de forma específica al resultado 4, de la meta 1 del pilar 1, que expresa:

“Se ha reducido hasta un 22 % las Necesidades Básicas Insatisfechas (NBI)³.”

El primer pilar (**Erradicación de la pobreza extrema**) está ampliamente vinculado con conceptos de bienestar, además, el NBI es uno de los pocos indicadores que pueden ser monitoreados tanto con los censos de población y las encuestas de hogares intercensales (Gomez, Alvarez, Lucarini, & Olmos, 1994). Este hace necesario el reflexionar sobre la necesidad de contar con un indicador que permita monitorear el bienestar de forma objetiva y se ajuste a los cambios en el tiempo de las variables que definen a un indicador.

¹Existen variaciones respecto a las variables, según la información que incluye un nuevo censo

²Ver Feres & Mancero, 2001, Pg.7

³Ministerio de Autonomías, 2013

1.3 Justificación

Es necesario definir un indicador alternativo al NBI, que supere sus limitaciones, respecto a la definición de sus variables y la ponderación de sus dimensiones. Que esté orientado principalmente a un criterio de información y variabilidad.

Montgomery, Gragnolati, Burke, & Paredes, 2000 propone que hay dos enfoques para calcular índices de bienestar; el primero, es el método directo el cual utiliza información sobre el ingreso, gasto y el consumo. El segundo método, es usar una medida **proxy** con los datos disponibles, destacando el autor la construcción de un índice alternativo de ingreso, gasto o consumo con información de bienes durables y características del hogar. Filmer, Pritchett, & Resources, 1998 utilizaron para el cálculo de un índice de bienestar, datos de bienes del hogar durables tales como; tenencia de televisor, refrigerador, auto, número de habitaciones en la vivienda, bicicleta, radio, procedencia del agua y otros servicios. Para esto utilizaron la metodología de Análisis de Componentes Principales (PCA).

Esta metodología ha sido ampliamente utilizada para los casos donde no hay disponibilidad de información explícita acerca del ingreso y el gasto. Harttgen & Vollmer, 2011 simulan el ingreso de los hogares, utilizando la metodología de Componentes Principales, para Bolivia, Burkina Faso, Indonesia y Zambia, utilizando datos de Encuestas de Salud e información macro económica. Con esto logran capturar la desigualdad intra y entre países a través del tiempo, subgrupos y características socio económicas.

1.4 Planteamiento del problema

El problema central es que no existe un indicador de bienestar asociado a los censos que este libre de subjetividad respecto a las variables que lo definen y las ponderaciones que se dan. La limitación principal de esta concepción es: se define de manera estática en el tiempo el concepto de bienestar.

Desde un punto de vista formal/matemático el NBI es construido como una combinación lineal del tipo $NBI_i = w_1 * d_{i1} + w_2 * d_{i2} + \dots$ donde las d_{ij} representa el valor de dimensión j para la vivienda i y los w_j son ponderaciones para cada dimensión, tanto las dimensiones y los ponderadores son seleccionados por un equipo técnico, según las características del país de estudio (Feres & Mancero, 2001), es decir, los criterios están condicionados a la información de los censos y a la definición del equipo técnico. A partir del trabajo se pretende proponer un modelo matemático capaz de definir variables (no dimensiones) y ponderaciones definidas por un criterio matemático y que solo sea limitado por las variables del censo y que permita medir el bienestar en un Censo.

1.5 Objetivos

1.5.1 Objetivo General

El objetivo general del trabajo es:

“ Definir un indicador dinámico, basado en el método de componentes principales para variables binarias, orientado a medir el bienestar de los hogares en los censos de población”

El concepto de dinámico⁴ dentro del trabajo está enfocado en 2 aspectos; (1) el primero, dado que se plantea que los componentes principales generan un indicador de bienestar basado en variables del censo, estas variables son distintas según el año del censo que se estudie, es decir, las variables no necesariamente son las mismas, (2) el segundo, dado que el conjunto de variables que se miden en un censo son similares entre censos, es posible aplicar los ponderadores dados por los componentes principales en un censo y aplicarlos para otro censo usando el conjunto de soporte común entre los censos. La utilidad de esta cualidad de los componentes principales, permitirá estudiar los cambios en composición y ponderación del indicador de bienestar durante el tiempo. Cualidad que puede emplearse también para explorar las diferencias en bienestar entre regiones dentro de un mismo país para un mismo tiempo.

1.5.2 Objetivos específicos

Los objetivos específicos son:

1. Desarrollar la teoría matemática asociada al indicador
2. Aplicar el indicador a los datos de los Censos Nacionales de Población y Vivienda 1992, 2001 y 2012

1.6 Operacionalización de los objetivos de estudio

La metodología en general sigue los siguientes pasos para su desarrollo:

1. **Revisión de la literatura / estado del arte:** Corresponde a una búsqueda en la literatura disponible sobre temas similares o avances en la misma línea de la investigación.
2. **Recopilación de información:** Se refiere a la gestión para conseguir las bases de datos necesarias para la aplicación de la metodología propuesta
3. **Tratamiento de la información:** En base a la revisión de la literatura y la información disponible se procede a adecuar las bases de datos para la aplicación del modelo definido.
4. **Estimación del modelo:** Una vez se cuente con la base lista se realizan las estimaciones, posterior a ello se verifica la sensibilidad del modelo.
5. **Ajuste del modelo:** En base al estudio de sensibilidad y dado algún imprevisto se realiza los ajustes que correspondan y se estima nuevamente el modelo.
6. **Análisis de los resultados:** Se generan los resultados a nivel de Municipios para los 3 censos (1992, 2001 y 2012) y se realiza un análisis de los resultados, comparando los cambios entre los censos y realizando un estudio comparado entre ellos.

⁴En Giorgio Alleva, 2016, pg.255 define a un indicador dinámico basado en la comparación de dos periodos de tiempo

7. **Conclusiones del trabajo:** Se procede a dar las conclusiones del trabajo en base a todo lo desarrollado

1.7 Alcances

Los alcances del trabajo en los diferentes aspectos son:

- **Temático:** El trabajo se desarrolla dentro del área de Estadística Multivariante, que corresponde al código **1209-09** del clasificador de la UNESCO
- **Temporal:** Los resultados del trabajo se limitan a los años censales, i.e. 1992, 2001 y 2012.
- **Geográfico:** La aplicación se realiza sobre los datos censales de Bolivia, sin embargo, es posible replicar la metodología para otros países. También es posible realizar estudios comparados entre países o periodos de tiempo

Capítulo 2

Marco Teórico

En éste capítulo se desarrollan los aspectos vinculados al propósito del trabajo, en 2.1 se aborda el marco normativo que relaciona la agenda patriótica, el plan nacional de desarrollo con la necesidad de medir el bienestar, en 2.2 se describe las características del método de Necesidades Básicas Insatisfechas (NBI), incluyendo sus limitaciones, en 2.3 se hace un recorrido cronológico de alternativas al NBI, basado en una revisión de literatura, en 2.4 se presenta algunas características de los métodos multivariados, en 2.5 se desarrolla la teoría de los componentes principales en este se explora los criterios para el elegir subconjuntos de componentes y variables, además se brinda un ejemplo sobre medidas de satisfacción. Finalmente, 2.6 se desarrolla los componentes principales para datos binarios.

2.1 Agenda Patriótica, Plan Nacional de Desarrollo y Objetivos de Desarrollo Sostenible

A la fecha Bolivia plantea como el primer pilar dentro del Plan de Desarrollo Económico y Social (PDES) 2016-2020 erradicar la pobreza extrema (Ministerio de Planificación del Desarrollo, 2015), el PDES se basa en lo que se denominó la agenda patriótica hacia el bicentenario de Bolivia (Ministerio de Autonomías, 2013) que dentro de su propuesta plantea 13 pilares cuyo objetivo es:

“Constituir la Bolivia Digna y Soberana, con el objetivo de levantar una sociedad y un Estado más incluyente, participativo, democrático, sin discriminación, racismo, odio, ni división.”

Estos 13 pilares son:

1. Erradicación de la pobreza extrema.
2. Socialización y universalización de los servicios básicos con soberanía para Vivir Bien.
3. Salud, educación y deporte para la formación de un ser humano integral.
4. Soberanía científica y tecnológica con identidad propia.
5. Soberanía comunitaria financiera sin servilismo al capitalismo financiero.

6. Soberanía productiva con diversificación y desarrollo integral sin la dictadura del mercado capitalista.
7. Soberanía sobre nuestros recursos naturales con nacionalización, industrialización y comercialización en armonía y equilibrio con la Madre Tierra.
8. Soberanía alimentaria a través de la construcción del saber alimentarse para Vivir Bien.
9. soberanía ambiental con desarrollo integral, respetando los derechos de la Madre Tierra.
10. Integración complementaria de los pueblos con soberanía.
11. Soberanía y transparencia en la gestión pública bajo los principios de no robar, no mentir y no ser flojo.
12. Disfrute y felicidad plena de nuestras fiestas, de nuestra música, nuestros ríos, nuestra selva, nuestras montañas, nuestros nevados, de nuestro aire limpio, de nuestros sueños.
13. Reencuentro soberano con nuestra alegría, felicidad, prosperidad y nuestro mar.

El objetivo asociado al primer pilar es de **erradicar la extrema pobreza en sus dimensiones material, social y espiritual**. Entendiendo como **pobreza material** la ausencia de acceso a servicios básicos y condiciones dignas de vida; **la pobreza social** se visibiliza en la predominancia del individualismo sobre los valores comunitarios; y la **pobreza espiritual** se expresa en la presencia de prácticas de consumismo, discriminación y racismo.

En el PDES se especifica con mas detalle el concepto y la forma de medir los resultados para el pilar, y los tipos de pobreza definidos, también, se plantean metas para cada uno de ellos, estos son:

- Erradicación de la pobreza material
 - **Meta 1:** Erradicación de la pobreza extrema material y reducción significativa de la pobreza moderada.
- Erradicación de la pobreza social, de toda forma de explotación, de la discriminación y del racismo
 - **Meta 2:** Combatir la pobreza social.
 - **Meta 3:** En Bolivia ya no existen las comunidades cautivas, ninguna forma de pongueaje y esclavismo, y explotación en el trabajo infantil.
- Erradicación de la pobreza espiritual y construcción del ser humano integral
 - **Meta 4:** Combatir la discriminación y el racismo.
 - **Meta 5:** Combatir la pobreza espiritual.
 - **Meta 6:** Construir un ser humano integral para Vivir Bien.

En particular los resultados para la meta 1 se presenta en el cuadro 2.1, esta meta esta directamente relacionada con indicadores de bienestar, entre ellos el indicador 3 explicita que se debe reducir un 22% en el indicador de necesidades básicas insatisfechas. A diferencia del resto de los indicadores planteados en la meta 1, el NBI es uno de los pocos indicadores que pueden ser monitoreados tanto con los censos de población y las encuestas de hogares intercensales (Gomez et al., 1994).

Cuadro 2.1: Resultados para la meta 1 del pilar: Erradicar la pobreza extrema

Resultado	Descripción
1	Se ha reducido al 9,5% la población en situación de pobreza extrema.
2	Se ha disminuido al 24% la población en situación de pobreza moderada.
3	Se ha reducido hasta un 22% las Necesidades Básicas Insatisfechas (NBI).
4	Se ha reducido hasta 25 veces la relación de ingresos entre el 10% más rico y el 10% más pobre.
5	Se ha cubierto al menos el 80% de los hogares más pobres y vulnerables con programas sociales.
6	Se ha ampliado el apoyo integral a niñas, niños y adolescentes en situación de calle, en centros de re-inserción social con la participación de los municipios.
7	Se ha promovido la incorporación de niñas y niños en situación de calle en familias sustitutas.
8	Se ha incrementado la cobertura de programas y servicios de Desarrollo Infantil Temprano (DIT) en niñas y niños menores de 4 años.
9	Se ha promovido el apoyo y atención integral a niñas, niños, jóvenes y adolescentes en situación de calle.
10	Se ha promovido el acceso de personas con discapacidad registradas en programas integrales de inclusión social basados en la comunidad
11	Se han impulsado programas de rehabilitación basados en la comunidad para la restitución y el ejercicio de los derechos de las personas con discapacidad
12	Se ha fortalecido el apoyo integral para personas adultas mayores en centros de acogida y otros espacios para su ejercicio al derecho a una vejez digna.

Fuente: Elaboración del autor en base a Ministerio de Planificación del Desarrollo, 2015

2.2 El Índice de Necesidades Básicas Insatisfechas (NBI) y sus limitaciones

El Índice de Necesidades Básicas Insatisfechas (NBI) fue introducido en la década de los 80 por la Comisión Económica para América Latina y el Caribe (Feres & Mancero, 2001), como un método directo para medir pobreza. El objetivo era aprovechar la información de los censos, demográficos y de vivienda, en la caracterización de la pobreza. Bajo este método, se elige una serie de indicadores censales que permiten constatar si los hogares satisfacen o no algunas de sus necesidades principales.

Dada la restricción del método a la información contenida en los censos, el tipo de necesidades que puede estudiar es limitado. Usualmente la insatisfacción de necesidades se evalúa en base a algunas características de la vivienda y ciertos rasgos demográficos del hogar. Así el concepto de pobreza implícito en el método NBI se limita en la práctica a unos pocas necesidades específicas, dejando de lado varios otros elementos relevantes del bienestar.

2.2.1 El método de NBI

Dado que la fuente principal para el método son los censos de población y vivienda, el primer paso para su definición es la selección de los indicadores, el proceso esta constituido, de manera general, por cuatro pasos:

1. Determinar el grupo de necesidades básicas mínimas susceptibles de estudiarse con la información del censo
2. Elegir indicadores censales que representen dichas necesidades
3. Definir el nivel crítico de satisfacción para cada necesidad (fijar la norma)
4. Asegurar que los indicadores seleccionados correspondan a situaciones de pobreza

Las necesidades consideradas suelen limitarse a:

- Acceso a una vivienda que asegure un estándar mínimo de habitabilidad para el hogar
- Acceso a servicios básicos que aseguren un nivel sanitario adecuado
- Acceso a educación básica
- Capacidad económica para alcanzar niveles mínimos de consumo

Aunque sería muy útil contar con información sobre el ingreso o el gasto de los hogares para medir su capacidad económica, los censos suelen no contener información sobre esas variables y cuando la contienen, ésta es poco fiable. El método NBI también debería considerar necesidades básicas como la salud y la nutrición, pero las variables respectivas no están disponibles en los censos de la región. Dadas estas limitaciones, el segundo paso en el proceso de selección de indicadores consiste en determinar cuáles son las dimensiones factibles de ser medidas dentro de cada necesidad básica, y qué variables censales serán utilizadas para dar cuenta de dichas dimensiones. Un ejemplo de estas se muestra en la figura 2.1.

Figura 2.1: Necesidades Básicas, dimensiones y variables censales

Necesidades Básicas	Dimensiones	VARIABLES CENSALES
Acceso a vivienda	a) Calidad de la vivienda	Materiales de construcción utilizados en piso, paredes y techo
	b) Hacinamiento	i) Número de personas en el hogar ii) Número de cuartos de la vivienda
Acceso a servicios sanitarios	a) Disponibilidad de agua potable	Fuente de abastecimiento de agua en la vivienda
	b) Tipo de sistema de eliminación de excretas	i) Disponibilidad de servicio sanitario ii) Sistema de eliminación de excretas
Acceso a educación	Asistencia de los niños en edad escolar a un establecimiento educativo	i) Edad de los miembros del hogar ii) Asistencia a un establecimiento educativo
Capacidad económica	Probabilidad de insuficiencia de ingresos del hogar	i) Edad de los miembros del hogar ii) Último nivel educativo aprobado iii) Número de personas en el hogar iv) Condición de actividad

Fuente: CEPAL / PNUD (1989)

Una vez definida las variables a emplear el siguiente paso es agregar la información, para ello se debe asignar puntajes específicos a cada posibilidad de satisfacción de una necesidad, de acuerdo con su grado de optimabilidad o precariedad. Es decir, para cada necesidad se establece un **indicador de logro**, cuyos valores muestran desde la forma mas extrema de carencia hasta la forma mas adecuada de satisfacer una necesidad. Por ejemplo para el caso de Bolivia en el censo de 2001 (Instituto Nacional de Estadística - Bolivia, 2005) se utiliza una escala que va desde cero hasta un valor máximo, determinado

por el numero de opciones de respuesta; en el caso de los materiales predominantes en el techo de una vivienda, los valores que se utilizan se muestran en el cuadro 2.2.

Cuadro 2.2: Calificación para el material predominante en el techo

Materiales predominantes	Calificación
Material de desecho	0
Paja, caña o palma	0.5
Calamina plancha	1
Tejas, losa de hormigón armado	1.5

Fuente: Elaboración del autor en base a Feres & Mancero, 2001

2.2.2 Críticas y limitaciones al método del NBI

Entre las fortalezas del NBI se tiene a a la efectiva utilización de la información censal y el bajo costo asociado a su medición, sin embargo, también existen criticas y limitaciones, entre ellas:

- Existe un alto costo en la periodicidad (cada 10 años)
- No es útil para medir pobreza reciente
- La comparabilidad inter temporal no sirve para evaluar la eficacia de una política
- Puede existir problemas en la comparabilidad geográfica, dado que se puede introducir un sesgo hacia la sub estimación de la pobreza urbana.
- Se tiene un problema de mensurabilidad: la posibilidad de ser clasificado como pobre no es igual para todos los hogares ya que ella depende de la estructura demográfica de los mismos (Gomez et al., 1994).

2.3 Alternativas al NBI

En las ultima década se empezó a plantear alternativas al NBI que superen sus limitaciones, Filmer et al., 1998 plantearon calcular un índice de activos (asset index) basado en componentes principales como sustituto al gasto en consumo del hogar para estimar la relación entre bienestar con la inscripción escolar en la India, defendiendo la posición de que este índice es mejor al consumo como medida de bienestar dada los diferentes errores de medición, este definió el índice de activos para cada hogar (A_j) como:

$$A_j = f_1 * \frac{(a_{j1} - a_1)}{s_1} + \dots + f_p * \frac{(a_{jp} - a_p)}{s_p} \quad (2.1)$$

Con f_1 el score para el primer activo determinado por el método de CPs, a_{j1} es el valor del primer activo en el j -esimo hogar, a_1 y s_1 son la media y la desviación estándar del primer activo considerado. Se definieron 21 activos para la construcción del índice.

Montgomery et al., 2000 plantea que uno de los mayores problemas de las encuestas (en especial las de demografía) es la falta de medición de indicadores de ingreso y consumo que son las que permiten entender el bienestar, para ello, define hacer un modelo que permita encontrar la relación entre el consumo y variables proxy. Realiza esto para 6 países; Ghana, Guatemala, Jamaica, Pakistan, Peru y Tanzania. De forma similar Lanjouw, Elbers, & Lanjouw, 2000 desarrolla un modelo de simulación para Ecuador, con el fin de construir mapas de pobreza monetaria en micro niveles usando información similar en censos y encuestas, estas variables están vinculadas a los activos del hogar.

En McKenzie, 2004 se introduce nuevamente un indicador basado en activos con la intención de medir la desigualdad, utiliza las encuestas de demografía y salud y desarrolla el método para México. Sugiere medir la desigualdad a nivel de comunidad con:

$$I_c = \frac{\sigma_c}{\sqrt{\lambda}} \quad (2.2)$$

Donde σ_c es la desviación estándar de y_i que corresponde a los valores del primer componente sobre los hogares de la comunidad, mientras que λ es el valor propio asociado al primer componente principal. Este indicador además cumple con los cuatro axiomas de la desigualdad; Anonimidad, escala independiente, independencia dada la población y la propiedad de Pigou-Dalton.

De manera mas reciente Harttgen & Vollmer, 2011 plantea una descomposición de la desigualdad sin datos de ingresos y consumo, usando un índice de activos basado en el primer componente principal que se utiliza para simular el ingreso a partir de estimar la distribución nacional del ingreso. Se utiliza las Encuestas de Demografía y Salud de 2003 de Bolivia, Burkina Faso, Indonesia y Zambia.

2.4 Métodos Multivariantes

El análisis multivariante es una bolsa mixta de métodos, es difícil establecer un esquema de clasificación de estas técnicas (Johnson & Wichern, 2007). Una clasificación distingue técnicas diseñadas a estudiar las relaciones de interdependencia de otras destinadas a estudiar las relaciones de dependencia. Otras clasificaciones se centran en el tipo de población y variables a estudiar. Mas abajo, se listan algunos problemas prácticos que permiten conectar el método con el problema de estudio. Estos problemas permiten apreciar la aplicación de las técnicas multivariantes en distintas áreas de estudio.

Los objetivos de diferentes investigadores, que de manera natural la estadística multivariante puede responder son las siguientes:

1. **Reducción de datos o simplificación estructural:** El fenómeno estudiado es representado de forma simple sin sacrificar información valiosa. Se espera que la interpretación sea sencilla
2. **Agrupación:** Grupos con similares objetos o variables son creados, basado en las características de los datos.
3. **Dependencia entre variables:** El interés es la relación natural de las variables

4. **Predicción:** La relación entre las variables pueden ser usadas para predecir los valores de una o mas variables sobre la observación de otras
5. **Construcción y prueba de hipótesis:** Test estadísticos son contrastados, formulados en términos de parámetros de poblaciones multivariantes. Esto puede ser usado para validar supuestos o reforzar las convicciones a priori.

Entre los tópicos más empleados dentro de la estadística multivariante se tiene:

- Correlaciones y regresiones, incluyendo; estructuras lineales y ecuaciones funcionales
- Análisis factorial y análisis de componentes principales
- Análisis de correlación canónica
- Análisis discriminante, clasificación y análisis cluster
- MANOVA, y los modelos lineales multivariante

2.4.1 La organización de los datos

Existen diversas formas de trabajar con los datos, en esta sección se introduce los conceptos necesarios para la organización de los datos en este trabajo.

Los datos multivariantes alrededor de un estudio sirven para comprender fenómenos sociales o físicos, a través de seleccionar $p \geq 1$ de variables o características. Los valores de estas variables son llamados registros para cada distinto item, individuo o unidad experimental. Se denota por x_{jk} un valor particular para la k -ésima variable que es observada en el item, o unidad j -ésima, esto es:

x_{jk} = Medida de la k esima variable sobre la unidad j esima

Consecuentemente, n medidas sobre p variables puede ser visto como:

	<i>Variable</i> ₁	<i>Variable</i> ₂	...	<i>Variable</i> _{k}	...	<i>Variable</i> _{p}
<i>Item</i> ₁ :	x_{11}	x_{12}	...	x_{1k}	...	x_{1p}
<i>Item</i> ₂ :	x_{21}	x_{22}	...	x_{2k}	...	x_{2p}
⋮	⋮	⋮	⋱	⋮	⋱	⋮
<i>Item</i> _{j} :	x_{j1}	x_{j2}	...	x_{jk}	...	x_{jp}
⋮	⋮	⋮	⋱	⋮	⋱	⋮
<i>Item</i> _{n} :	x_{n1}	x_{n2}	...	x_{nk}	...	x_{np}

Este puede ser definido en forma de una matriz rectangular, llamada \mathbf{X} , de n filas y p columnas:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2k} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & x_{j2} & \dots & x_{jk} & \dots & x_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} & \dots & x_{np} \end{bmatrix}$$

2.4.2 Álgebra de Matrices y Vectores aleatorios

En este apartado se desarrollan los conceptos necesarios dentro del álgebra lineal, vinculados al método de componentes principales que se desarrolla en la sección siguiente, la notación y definiciones son adaptaciones a partir de Johnson & Wichern, 2007.

Vectores

Un arreglo \mathbf{x} de n números reales x_1, x_2, \dots, x_n es llamado un vector y la notación de este es:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{ó} \quad \mathbf{x}' = [x_1, x_2, \dots, x_n] \quad (2.3)$$

Donde el símbolo $'$ denota la operación de la transpuesta, de columna a fila.

El vector \mathbf{x} puede ser representado geoméricamente como una línea recta desde el origen hacia las direcciones definidas en cada una de las n dimensiones, x_1, x_2, \dots, x_n .

La longitud de un vector \mathbf{x} con n elementos es definida como:

$$L_x = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\mathbf{x}'\mathbf{x}} \quad (2.4)$$

Donde $\mathbf{x}'\mathbf{x}$ es el producto interior de dos vectores.

Otro concepto importante es el ángulo entre vectores, sean los vectores $\mathbf{x}' = [x_1, x_2, \dots, x_n]$ y $\mathbf{y}' = [y_1, y_2, \dots, y_n]$ de n elementos, el ángulo θ entre estos dos vectores esta dado por:

$$\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_x L_y} \quad (2.5)$$

Notar que $\cos(\theta) = 0$ solo si $\mathbf{x}'\mathbf{y} = 0$, en esos casos se dice que \mathbf{x} e \mathbf{y} son perpendiculares.

Se dice que un set de vectores $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ son linealmente dependientes si existen constantes c_1, c_2, \dots, c_k no todas cero (0) tal que:

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k = \mathbf{0} \quad (2.6)$$

La dependencia lineal implica que al menos uno de los vectores en el set de vectores puede ser escrito como una combinación de los otros vectores. Vectores de la misma dimensión n que no son linealmente dependientes, se dice que son linealmente independientes.

La proyección (o sombra) de un vector \mathbf{x} sobre un vector \mathbf{y} esta dada por:

$$\text{Proyección de } \mathbf{x} \text{ sobre } \mathbf{y} = \frac{(\mathbf{x}'\mathbf{y})}{\mathbf{y}'\mathbf{y}} \mathbf{y} = \frac{(\mathbf{x}'\mathbf{y})}{L_y} \frac{1}{L_y} \mathbf{y} \quad (2.7)$$

Donde la longitud del vector $L_y^{-1}\mathbf{y}$ es igual a la unidad, la longitud de la proyección es:

$$\text{Longitud de la proyección} = \frac{|\mathbf{x}'\mathbf{y}|}{L_y} = L_x \left| \frac{\mathbf{x}'\mathbf{y}}{L_x L_y} \right| = L_x |\cos(\theta)| \quad (2.8)$$

Matrices

Una matriz es un arreglo rectangular de números reales, se denota un arreglo de n filas y p columnas como:

$$\mathbf{A}_{(n \times p)} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{bmatrix}$$

La operación transpuesta de una matriz se denota por \mathbf{A}' , esta cambia las columnas en filas, de tal forma que la primera columna de \mathbf{A} se convierte la primera fila de \mathbf{A}' .

Se definen las siguientes operaciones para las matrices:

- **Multiplicación por una constante c :** $c\mathbf{A} = \{c * a_{ij}\}$
- **Suma:** $\mathbf{A} + \mathbf{B} = \{a_{ij} + b_{ij}\}$
- **Producto de matrices:** $\mathbf{A}_{(n,k)} * \mathbf{B}_{(k,p)} = \mathbf{AB}_{(n,p)} = \{\text{Entrada (ij)} \sum_{l=1}^k a_{il}b_{lj}\}$

Sea \mathbf{I} una matriz cuadrada con unos en la diagonal y 0 en el resto de las posiciones. Un resultado a partir de $\mathbf{I}_{(n,n)}$ es que para una matriz $\mathbf{A}_{(n,n)}$, $\mathbf{A} = \mathbf{IA} = \mathbf{AI}$. La matriz \mathbf{I} se denomina la matriz identidad.

Si $\mathbf{A} = \mathbf{A}'$ la matriz \mathbf{A} se denomina una matriz simétrica, es decir $a_{ij} = a_{ji}$ para todo i, j .

Si existe una matriz \mathbf{B} tal que:

$$\mathbf{B}_{(k,k)} \mathbf{A}_{(k,k)} = \mathbf{A}_{(k,k)} \mathbf{B}_{(k,k)} = \mathbf{I}_{(k,k)} \quad (2.9)$$

Entonces \mathbf{B} es llamada la inversa de \mathbf{A} y esta se denota por \mathbf{A}^{-1}

La condición para la existencia de \mathbf{A}^{-1} es que las k columnas $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ de \mathbf{A} son linealmente independientes. Esto implica que la existencia de \mathbf{A}^{-1} es equivalente a:

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k = \mathbf{0} \text{ Solo si } c_1 = c_2 = \dots = c_k \quad (2.10)$$

Una matriz cuadrada especial es la que cumple:

$$\mathbf{Q}\mathbf{Q}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I} \text{ ó } \mathbf{Q}' = \mathbf{Q}^{-1} \quad (2.11)$$

\mathbf{Q} se denomina matriz ortogonal. El nombre deriva de la propiedad que si \mathbf{Q} tiene una i -ésima fila \mathbf{q}'_i , entonces $\mathbf{Q}\mathbf{Q}' = \mathbf{I}$ implica que $\mathbf{q}'_i\mathbf{q}_i = 1$ y $\mathbf{q}'_i\mathbf{q}_j = 0$ para $i \neq j$, de tal forma que las filas tienen una longitud de 1 y son mutuamente perpendiculares

(ortogonales). Las columnas de $\mathbf{Q}'\mathbf{Q}$ tienen las mismas propiedades.

A continuación se describe un concepto fundamental dentro del álgebra de matrices para la estadística multivariante. Sea \mathbf{A} una matriz cuadrada, se dice que esta tiene un *eigenvalor* λ , con su correspondiente *eigenvector* $\mathbf{x} \neq 0$, si:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (2.12)$$

Normalmente \mathbf{x} es un vector normalizado, es decir $\mathbf{x}'\mathbf{x} = 1$, es conveniente denotar a un vector normalizado como \mathbf{e} , a partir de ello se establece el siguiente resultado:

Sea \mathbf{A} una matriz simétrica cuadrada de dimension k, k . Entonces \mathbf{A} tiene k pares de *eigenvalores* y *eigenvectores*, nombrados:

$$\lambda_1, \mathbf{e}_1 \quad \lambda_2, \mathbf{e}_2 \quad \dots \quad \lambda_k, \mathbf{e}_k \quad (2.13)$$

Los *eigenvectores* son elegidos tal que satisfacen $1 = \mathbf{e}'_1\mathbf{e}_1 = \dots = \mathbf{e}'_k\mathbf{e}_k$ y son mutuamente perpendiculares $0 = \mathbf{e}'_i\mathbf{e}_j$ para todo $i \neq j$. Los *eigenvalores* son únicos a menos que dos o más *eigenvectores* sean iguales.

Descomposición espectral y matriz definida positiva

La descomposición espectral de una matriz simétrica \mathbf{A} de dimension k, k es dada por:

$$\mathbf{A} = \lambda_1 \underset{(k,k)}{\mathbf{e}_1} \underset{(k,1)(1,k)}{\mathbf{e}'_1} + \lambda_2 \underset{(k,1)(1,k)}{\mathbf{e}_2} \underset{(k,1)(1,k)}{\mathbf{e}'_2} + \dots + \lambda_k \underset{(k,1)(1,k)}{\mathbf{e}_k} \underset{(k,1)(1,k)}{\mathbf{e}'_k} \quad (2.14)$$

Donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son los *eigenvalores* de \mathbf{A} y $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ son los *eigenvectores* normalizados. Tal que $\mathbf{e}'_i\mathbf{e}_i = 1$ y $\mathbf{e}'_i\mathbf{e}_j = 0$ para $i = 1, 2, \dots, k, i \neq j$.

Una matriz \mathbf{A} es definida positiva si:

$$0 < \mathbf{x}'\mathbf{A}\mathbf{x} \quad (2.15)$$

Para todo vector $\mathbf{x} \neq 0$. Por que $\mathbf{x}'\mathbf{A}\mathbf{x}$ cuenta con solo términos al cuadrado x_i^2 y términos $x_i x_k$, esta se llama la forma cuadrática.

Sea \mathbf{A} de dimension k, k , una matriz simétrica y definida positiva con la descomposición espectral $\mathbf{A} = \sum_{i=1}^k \lambda_i \mathbf{e}_i \mathbf{e}'_i$. Sea la matriz \mathbf{P} una matriz compuesta de los *eigenvectores* en sus columnas $\mathbf{P} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]$. Entonces:

$$\mathbf{A} = \sum_{i=1}^k \lambda_i \underset{(k,1)(1,k)}{\mathbf{e}_i} \underset{(k,1)(1,k)}{\mathbf{e}'_i} = \underset{(k,k)}{\mathbf{P}} \underset{(k,k)}{\mathbf{\Lambda}} \underset{(k,k)}{\mathbf{P}'} \quad (2.16)$$

Donde $\mathbf{P}\mathbf{P}' = \mathbf{P}'\mathbf{P} = \mathbf{I}$ y $\mathbf{\Lambda}$ es una matriz diagonal con los *eigenvalores* en su diagonal, $diagonal(\mathbf{\Lambda}) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$

Vectores y matrices aleatorias

Un vector aleatorio es un vector que contiene variables aleatorias. De manera similar una matriz aleatoria es una matriz que contiene variables aleatorias.

Suponer que $\mathbf{X} = [X_1, X_2, \dots, X_p]$ es un vector aleatorio de $p, 1$. Cada elemento de \mathbf{X} es una variable aleatoria con su propia distribución de probabilidad. Las medias marginales μ_i y la varianza σ_i^2 son definidas como $\mu_i = E(X_i)$ y $\sigma_i^2 = E(X_i - \mu_i)^2$, $i = 1, 2, \dots, p$. La covarianza entre dos variables aleatorias X_i y X_j se denota como $\sigma_{ij} = E(X_i - \mu_i)(X_j - \mu_j)$, cuando $i = j$ se obtiene la varianza $\sigma_i^2 = \sigma_{ii}$.

La media y la covarianza del vector aleatorio \mathbf{X} de dimensión $p, 1$ puede ser dispuesto en forma matricial. El valor esperado de cada elemento es contenido en el vector de medias $\mu = E(\mathbf{X})$, y las p varianzas σ_{ii} y las $p(p-1)/2$ distintas covarianzas $\sigma_{ik} (i < k)$ son contenidas en la matriz simétrica de varianzas y covarianzas $\Sigma = E(\mathbf{X} - \mu)(\mathbf{X} - \mu)'$. Esto es:

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \mu_p \end{bmatrix} = \mu \quad (2.17)$$

y,

$$\begin{aligned} \Sigma &= E(\mathbf{X} - \mu)(\mathbf{X} - \mu) \\ &= E \left(\begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots X_p - \mu_p \end{bmatrix} [X_1 - \mu_1, X_2 - \mu_2, \dots, X_p - \mu_p] \right) \\ &= \begin{bmatrix} E(X_1 - \mu_1)^2 & E(X_1 - \mu_1)(X_2 - \mu_2) & \dots & E(X_1 - \mu_1)(X_p - \mu_p) \\ E(X_2 - \mu_2)(X_1 - \mu_1) & E(X_2 - \mu_2)^2 & \dots & E(X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p - \mu_p)(X_1 - \mu_1) & E(X_p - \mu_p)(X_2 - \mu_2) & \dots & E(X_p - \mu_p)^2 \end{bmatrix} \end{aligned}$$

ó,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} \quad (2.18)$$

2.5 Método de componentes principales

El método de Análisis de Componentes Principales se ocupa de explicar la estructura de varianza y covarianza de un grupo de variables a través de unas pocas combinaciones lineales de este grupo de variables. En general sus objetivos son (1) la reducción de los datos y (2) la interpretación. (Johnson & Wichern, 2007).

Primer componente principal = Combinacion lineal $a'_1 X$ que maximiza $Var(a'_1 X)$ sujeto a $a'_1 a_1 = 1$
 Segundo componente principal = Combinacion lineal $a'_2 X$ que maximiza $Var(a'_2 X)$ sujeto a $a'_2 a_2 = 1$ y $Cov(a'_1 X, a'_2 X) = 0$

Para el i – esimo paso:

i – esimo componente principal = Combinacion lineal $a'_i X$ que maximiza $Var(a'_i X)$ sujeto a $a'_i a_i = 1$ y $Cov(a'_i X, a'_k X) = 0$ para $k < i$

Resultado 2.1 Sea Σ la matriz de covarianza asociada al vector aleatorio $\mathbf{X}' = [X_1, X_2, \dots, X_p]$. Sea Σ matriz con pares de valores propios y vectores propios $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ donde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Entonces el i – esimo componente principal esta dado por:

$$Y_i = e'_i \mathbf{X} = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p \quad (2.25)$$

Con esta elección:

$$\begin{aligned} Var(Y_i) &= e'_i \Sigma e_i = \lambda_i, & i = 1, 2, \dots, p \\ Cov(Y_i, Y_k) &= e'_i \Sigma e_k = 0 & i \neq k \end{aligned} \quad (2.26)$$

Si algunos λ_i son iguales, la elección de los coeficientes del vector e_i , por lo tanto Y_i , no son únicos.

Demostración 2.1 En base a la Maximización de la forma cuadrática para puntos sobre la esfera unitaria¹ (MCEU)², con $\mathbf{B} = \Sigma$, tal que:

$$\max_{a \neq 0} \frac{a' \Sigma a}{a' a} = \lambda_1 \quad (\text{Alcanzado cuando } a = e_1)$$

Pero $e'_1 e_1 = 1$, ya que los vectores propios son normalizados. Así,

$$\max_{a \neq 0} \frac{a' \Sigma a}{a' a} = \lambda_1 = \frac{e'_1 \Sigma e_1}{e'_1 e_1} = e'_1 \Sigma e_1 = Var(Y_1)$$

De manera similar, usando MCEU, se tiene:

$$\max_{a \perp e_1, e_2, \dots, e_k} \frac{a' \Sigma a}{a' a} = \lambda_{k+1} \quad k = 1, 2, \dots, p-1$$

Para la elección $a = e_{k+1}$ con $e'_{k+1} e_i = 0$, para $i = 1, 2, \dots, k$ y $k = 1, 2, \dots, p-1$

¹La esfera unitaria se define como $S = \{x \in \mathbf{R}^p : x'x = 1\}$

²Ver página 83, Johnson & Wichern, 2007

$$\frac{e'_{k+1} \Sigma e_{k+1}}{e'_{k+1} e_{k+1}} = e'_{k+1} \Sigma e_{k+1} = \text{Var}(Y_{k+1})$$

Pero $e'_{k+1}(\Sigma e_k + 1) = e'_{k+1}(\lambda_{k+1} e_k + 1) = \lambda_{k+1}$, así, $\text{Var}(Y_{k+1} = \lambda_{k+1})$. Queda por mostrar que e_i perpendicular a e_k (esto es, $e'_i e_k = 0$, $i \neq k$) da $\text{Cov}(Y_i, Y_k) = 0$. Ahora los vectores propios de Σ son ortogonales si todos los valores propios $\lambda_1, \lambda_2, \dots, \lambda_p$ son distintos. Si los vectores propios no son todos distintos, los vectores propios correspondientes a valores propios comunes pueden elegirse para llegar a ser ortogonales. Por lo tanto, para cualquier vector propio e_i y e_k , $e'_i e_k = 0$, $i \neq k$. Ya que $\Sigma e_k = \lambda_k e_k$, pre multiplicando por e'_k se tiene:

$$\text{Cov}(Y_i, Y_k) = e'_i \Sigma e_k = e'_i \lambda_k e_k = \lambda_k (e'_i e_k) = 0$$

Para cualquier $i \neq k$, y la demostración está completa

Del resultado 2.1, los componentes principales son in correlacionados y tienen varianzas iguales a los valores propios de Σ

Resultado 2.2 Sea $\mathbf{X}' = [X_1, X_2, \dots, X_p]$ con matriz de covarianza Σ , con pares de valores propios y vectores propios $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ donde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Sea $Y_1 = e'_1 \mathbf{X}$, $Y_2 = e'_2 \mathbf{X}$, \dots , $Y_p = e'_p \mathbf{X}$ los componentes principales. Entonces:

$$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \sum_{i=1}^p \text{Var}(X_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p = \sum_{i=1}^p \text{Var}(Y_i)$$

Demostración 2.2 Se sabe que $\text{tr}(\Sigma) = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp}$. También la descomposición espectral de $\Sigma = P \Lambda P'$ donde Λ es matriz diagonal con los valores propios de Σ y $P = [e_1, e_2, \dots, e_p]$ tal que $PP' = P'P = I$. Por propiedades de la traza se tiene:

$$\text{tr}(\Sigma) = \text{tr}(P \Lambda P') = \text{tr}(\Lambda P' P) = \text{tr}(\Lambda I) = \text{tr}(\Lambda) = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp}$$

Así,

$$\sum_{i=1}^p \text{Var}(X_i) = \text{tr}(\Sigma) = \text{tr}(\Lambda) = \sum_{i=1}^p \text{Var}(Y_i)$$

El resultado 2.2 implica que:

$$\begin{aligned} \text{Varianza total de la población} &= \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} \\ &= \lambda_1 + \lambda_2 + \dots + \lambda_p \end{aligned} \quad (2.27)$$

En consecuencia, la proporción de la varianza total explicada por el k -ésimo componente principal (PVT_k) es:

$$PVT_k = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, \quad k = 1, 2, \dots, p \quad (2.28)$$

Los coeficientes presentes en cada eigenvector son $e'_i = [e_{i1}, \dots, e_{ik}, \dots, e_{ip}]$. La magnitud e_{ik} mide la importancia (**contribución**) de la k -ésima variable al i -ésimo componente principal, independientemente de las otras variables. En particular, e_{ik} es proporcional al coeficiente de correlación entre Y_i y X_k .

Resultado 2.3 Si $Y_1 = e'_1 \mathbf{X}$, $Y_2 = e'_2 \mathbf{X}$, \dots , $Y_p = e'_p \mathbf{X}$ son los componentes principales obtenidos de la matriz de covarianza Σ , entonces:

$$\rho_{Y_i, X_k} = \frac{e_{ik} \sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \quad i, k = 1, 2, \dots, p \quad (2.29)$$

Son las correlaciones entre el componente principal Y_i y la variable X_k .

Demostración 2.3 Sea $a'_k = [0, \dots, 0, 1, 0, \dots, 0]$ tal que $X_k = a'_k \mathbf{X}$ y $Cov(X_k, Y_i) = Cov(a'_k \mathbf{X}, e'_i \mathbf{X}) = a'_k \Sigma e_i$. Dado que $\Sigma e_i = \lambda_i e_i$, $Cov(X_k, Y_i) = a'_k \lambda_i e_i = \lambda_i e_{ik}$. Como $Var(Y_i) = \lambda_i$ (2.26) y $Var(X_k) = \sigma_{kk}$, se tiene:

$$\rho_{Y_i, X_k} = \frac{Cov(X_k, Y_i)}{\sqrt{Var(Y_i)} \sqrt{Var(X_k)}} = \frac{\lambda_i e_{ik}}{\sqrt{\lambda_i} \sqrt{\sigma_{kk}}} = \frac{e_{ik} \sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \quad i, k = 1, 2, \dots, p$$

2.5.1 Eligiendo un sub conjunto de componentes principales

Una pregunta común en componentes principales es respecto a elegir la cantidad adecuada de componentes principales para el total de varianza en \mathbf{X} , el principal objetivo en componentes es de remplazar los p elementos en \mathbf{X} , por un número mucho más pequeño m número de componentes principales, es crucial conocer cuán pequeño puede ser m sin tener el riesgo de perder información seria.

A continuación se desarrollan reglas para la elección de los m componentes, las primeras tres son reglas ad hoc (rule-of-thumb), funcionan en la práctica y son sencillas de comprender. En la última se presenta una regla estadística, ésta no requiere supuestos sobre la distribución.

Método 1: Porcentaje acumulado de la varianza total

Es uno de los métodos más empleados de manera informal, la idea es seleccionar en base al porcentaje acumulado de la varianza total de los componentes principales, normalmente se seleccionan los que logran entre un 80 % y 90 %. El número de componentes principales es entonces el valor más pequeño de m para el que el porcentaje elegido es superado. Una definición de este porcentaje acumulado se muestra en 2.30.

$$t_m = 100 * \frac{\sum_{k=1}^m \lambda_k}{\sum_{k=1}^p \sigma_{kk}} \quad (2.30)$$

Eligiendo un t^* entre 70 % y 90 % se retiene los componentes principales donde m es el entero más pequeño tal que $t_m > t^*$. El mejor valor para t^* generalmente es más pequeño mientras p o n crecen. El rango entre 70 % y 90 % puede llegar a veces a ser bajo o alto, esto depende de la particularidad de los datos que se observen (I. T. Jolliffe, 2002). Por ejemplo, un valor de 90 % es adecuado cuando uno o dos componentes se muestran dominantes de la fuente de variación. Cuando p es muy grande, la elección de m correspondiente al 70 % puede dar un valor alto en m que normalmente es poco práctico. En estos casos el porcentaje debiera ser más bajo.

Método 2: Tamaño de la varianza de los componentes principales (Kaiser's Rule)

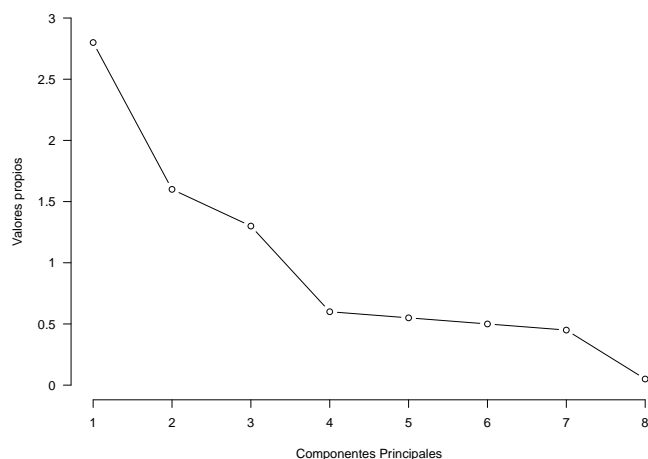
Esta regla es construida específicamente para el uso de componentes principales con la matriz de correlación, ésta puede ser adaptada para algunos tipos de matrices de covarianza. La idea detrás de la regla es que si todos los elementos de \mathbf{X} son independientes, entonces los componentes son iguales a las variables originales y todas tienen varianza unitaria. Así, cualquier componente principal con una varianza inferior a 1 contiene menos información que la variable original, por lo tanto no debe ser retenida.

Esta regla puede ser adaptada para matrices de covarianza, poniendo como punto de corte λ^* a el promedio de los valores propios $\bar{\lambda}$ o mejor, un punto de corte más bajo como $\lambda^* = 0,7\bar{\lambda}$. Esta regla es inadecuada cuando las varianzas en la matriz de covarianzas son muy diferentes, normalmente esto sucede cuando las unidades de \mathbf{X} son distintas, una regla sugerida para estos casos puede verse en I. T. Jolliffe, 2002.

Método 3: El gráfico del Scree plot y el plot Log-eigenvalue

Las dos primeras reglas involucran un grado de subjetividad en la elección del punto de corte basado en t^* y λ^* . El scree plot que fue introducido y nombrado por Cattell, 1966 es otro método subjetivo, este involucra buscar en un plot de los λ_k y decidir que valor k retener en base a la pendiente o un quiebre visual definido como un **codó** (elbow) en la figura, un ejemplo se muestra en la figura 2.2.

Figura 2.2: Ejemplo de un scree plot para datos simulados



Una alternativa al scree plot que fue desarrollada en la ciencia atmosférica, la idea es hacer el plot de $\log(\lambda_k)$ con k , el método es conocido como el diagrama log-eigenvalue (LEV). El procedimiento se basa en elegir los componentes previos a que el gráfico se convierta aproximadamente en una línea recta.

Método 4: Correlación parcial

Para los componentes principales basados en la matriz de correlaciones, Velicer, 1976 sugirió que las correlaciones parciales entre las p variables, da el valor de los m componentes a retener. El criterio propuesto es la promedio de las correlaciones parciales al cuadrado.

$$V = \sum_{i=1, i \neq j}^p \sum_{j=1}^p \frac{(r_{ij}^*)^2}{p(p-1)} \quad (2.31)$$

Donde r_{ij}^* es la correlación parcial entre las variables i y j dado los primeros m CPs. El estadístico r_{ij}^* es definido como la correlación entre los residuos de la regresión lineal de la i –ésima variable sobre los primeros m CPs, y los residuos de la regresión correspondiente de la j –ésima variable sobre los m CPs. Esto por lo tanto mide la relación lineal estrecha entre la i –ésima y j –ésima variable después de remover el efecto común sobre los primeros m CPs.

El criterio V disminuye primero y luego aumenta a medida que m aumenta y Velicer, 1976 sugiere que el valor óptimo de m corresponde al valor mínimo del criterio.

2.5.2 Eligiendo un sub conjunto variables

Cuando p , el número de variables observadas es muy grande, a veces sucede el caso que un sub conjunto de m variables, con $m < p$, contiene virtualmente toda la información disponible en todas las p variables. Por ello es útil determinar de forma apropiada el valor m y decidir que grupo de variables m son las mejores. La solución a estos dos problemas, la elección de m y la selección del mejor grupo de variables, depende del propósito con el cual las variables se han reunido. Si el propósito es simplemente preservar la mayor cantidad de varianza en \mathbf{X} , entonces los CPs pueden ser utilizados ampliamente para resolver ambos problemas. En este apartado se desarrolla el problema de encontrar un subgrupo de \mathbf{X} que represente de la mejor forma posible la variación interna de \mathbf{X} .

En la sección 2.5.1 se discutieron técnicas que permiten encontrar el número de CPs que mejor representan la variabilidad de \mathbf{X} , ello también puede ser interpretado como encontrar la dimensión efectiva de \mathbf{X} . Si \mathbf{X} puede ser descrito con solo m componentes principales, entonces, \mathbf{X} puede ser remplazado por un sub grupo de m variables, con una relativa pequeña pérdida de información. En I. Jolliffe, 2013 se discuten varios métodos para elegir m del tal forma que \mathbf{X} preserve en mayor medida su variación, algunos métodos están basados en componentes principales, otros están basados en análisis cluster. Se describen tres métodos que utilizan CPs.

1. Asociar una variable con cada una de los últimos $m_1^*(= p - m_1)$ CPs y eliminar a las m_1^* variables. Esto puede ser hecho solo una vez o de forma iterativa. En el último caso, se calcula nuevamente el CPs sobre las m_1 variables restantes y un nuevo set de m_2^* es eliminado. Un tercer CPs puede ser estimado y repetir el mismo procedimiento sobre los $p - m_1^* - m_2^*$ variables. El procedimiento sigue hasta que se considera que ya no es necesario eliminar variables. La elección de los m_1^*, m_2^*, \dots está basado sobre un criterio determinado por el tamaño de los valores propios λ_k

El razonamiento detrás de este método es que pequeños valores propios corresponden a relaciones casi constantes entre un sub conjunto de variables. Si una de esas variables es eliminada, se pierde poca información.

2. Asociar un set de m^* variables en bloque con los últimos m^* CPs, entonces eliminar estas variables (I. Jolliffe, 2013). Este tipo de método, con las m^* variables también elegidas para maximizar la suma de cuadrados de los coeficiente en las últimas m^* CPs o para ser aquellas variables m^* que son las mejores predictoras para una regresión sobre los primeros $m = p - m^*$ CPs. La elección de m^* es nuevamente basada en el tamaño de las λ_k .
3. Asociar una variable con cada una de las primeros CPs, es decir, la variable que ya no se haya elegido con el coeficiente más alto en valor absoluto en cada CPs sucesiva. Estas m variables se conservan, y el resto $m^* = p - m$ se eliminan.

2.5.3 Un ejemplo de CPs sobre medidas de satisfacción

Para mostrar la aplicación del método de CPs, se presenta un ejemplo clásico acerca de medidas de satisfacción relacionadas a las condiciones laborales del personal de las fuerzas armadas en Estados Unidos, este ejemplo se puede encontrar con más detalle en Dunteman, 1989. Las variables analizadas fueron parte de un set más grande de variables recolectadas. Estas son; Satisfacción con el trabajo (SJ), Satisfacción con la capacitación en el trabajo (SJT), Satisfacción con las condiciones de trabajo (SWC), Satisfacción con el seguro médico (SMC) Y Satisfacción con el seguro dental (SDC). Cada respuesta fue definida en una escala de satisfacción de 1 (Muy insatisfecho) a 5 (Muy Satisfecho) para cada una de las satisfacciones. La matriz de correlaciones se encuentra en el cuadro 2.3 que fue generada en base a una sub muestra de $N = 9147$ casos de militares casados. Notar que todas las correlaciones son positivas y varían entre 0,162 (la mas baja) y 0,620 (la mas alta).

Cuadro 2.3: Matriz de correlación (ρ) para las 5 variables de satisfacción, $N = 9147$

Variabes		(SJ= X_1)	(SJT= X_2)	(SWC= X_3)	(SMC= X_4)	(SDC= X_5)
Satisfacción con el trabajo	(SJ= X_1)	1	0.451	0.511	0.197	0.162
Satisfacción con la capacitación en el trabajo	(SJT= X_2)		1	0.445	0.252	0.238
Satisfacción con las condiciones de trabajo	(SWC= X_3)			1	0.301	0.227
Satisfacción con el seguro medico	(SMC= X_4)				1	0.62
Satisfacción con el seguro dental	(SDC= X_5)					1

Fuente: Elaboración del autor en base a Dunteman, 1989

Dada las cinco variables de satisfacción es posible extraer 5 componentes principales de la matriz de correlación, el cuadro 2.4 muestra los valores y vectores propios del ejemplo y también el PVT .

Los elementos en la primera columna son los ponderadores asociados con la composición lineal que maximiza la varianza y está asociado al primer componente principal, este componente es $0,442SJ + 0,457SJT + 0,479SWC + 0,443SMC + 0,412SDC$. Los ponderadores son casi iguales para cada una de las variables de satisfacción, el primer componente podría interpretarse como una medida general de satisfacción. La varianza del primer componente principal es 2.370, este explica el 47.4 % de la varianza total de las cinco variables. El segundo componente define la relación $0,443SJ + 0,29SJT +$

Cuadro 2.4: Eigenvalores y eigenvectores de la matriz de correlación de las variables de satisfacción y el *PVT*

Variable	Eigenvectores				
	e_1	e_2	e_3	e_4	e_5
(SJ)	0.442	0.443	0.301	-0.716	0.074
(SJT)	0.457	0.29	-0.832	0.114	0.034
(SWC)	0.479	0.308	0.454	0.658	-0.185
(SMC)	0.443	-0.531	0.095	0.06	0.714
(SDC)	0.412	-0.586	0.032	-0.191	-0.67
λ_i	2.370	1.202	0.573	0.484	0.373
PVT_i	0.474	0.240	0.115	0.097	0.075
Acumulado PVT_i	0.474	0.714	0.829	0.926	1.000

Fuente: Elaboración del autor en base a Dunteman, 1989

$0,308SWC - 0,531SMC - 0,586SDC$, los tres primeros ponderadores son positivos y están asociados a la satisfacción laboral y los siguientes negativos y están asociados a la satisfacción en temas de seguro médico, el segundo componentes puede ser interpretado como un contraste entre satisfacción laboral y satisfacción en la salud, altos niveles en este componente esta asociado con una fuerte satisfacción laboral y bajos niveles están asociados a una baja satisfacción en términos del seguro medico. El segundo componente tiene una varianza de 1,202 que explica el 24 % de la varianza. Juntos los dos componentes explican el 71 % de la varianza en las 5 variables de satisfacción, el resto de los componentes no pueden ser interpretados dado que no aportan información significativa dada el porcentaje de varianza que explican, así, las 5 variables planteadas son reducidas solo a 2.

Las cargas factoriales o correlación entre los componentes y las variables se presentan en el cuadro 2.5, muchos autores prefieren interpretar el patrón de las correlaciones respecto a los ponderadores mostrados en el cuadro 2.4. El tamaño de la correlación para un componente principal refleja de forma directa la importancia del componente en la explicación de la varianza de las variables originales. Por ello es importante notar que los últimos componentes principales presentan correlaciones cercanas a 0 con la mayoría de las variables originales.

Cuadro 2.5: Matriz de cargas factoriales (correlaciones *CPs* y variables) para las variables de satisfacción

Variable	Componentes principales				
	1	2	3	4	5
(SJ)	0.680	0.485	0.228	-0.498	0.045
(SJT)	0.704	0.318	-0.630	0.079	0.021
(SWC)	0.738	0.338	0.344	0.458	-0.113
(SMC)	0.682	-0.582	0.072	0.042	0.436
(SDC)	0.634	-0.642	-0.024	-0.133	-0.409

Fuente: Elaboración del autor en base a Dunteman, 1989

2.6 Componentes Principales para datos Binarios

Cuando se usa los componentes principales como técnica descriptiva, no hay razón para que las variables en el análisis sean de un tipo en particular. En un extremo la matriz de datos X puede tener una distribución normal multivariada, en cuyo caso se pueden usar todos los resultados inferenciales relevantes. En el extremo opuesto, las variables podrían ser una mezcla de variables continuas, ordinales o incluso binarias (0/1). Es cierto que las varianzas, las covarianzas y las correlaciones tienen una relevancia especial para la distribución Normal X , y que las funciones lineales de las variables binarias son menos fácil de interpretar que las funciones lineales de las variables continuas. “Sin embargo, el objetivo básico de PCA (resumir la mayor parte de la *variación* que está presente en el conjunto original de variables p utilizando un número menor de variables derivadas) se puede lograr independientemente de la naturaleza de las variables originales”... (I. Jolliffe, 2013).

Autores como Gower, 1966 y Freeman & Jackson, 1992 justifican el uso de componentes principales para variables binarias, en el caso de Gower establece en el apartado 4.1

“Aunque un análisis convencional de componentes principales de $(0, 1)$ datos puede parecer dudoso, lo anterior muestra que es exactamente equivalente a suponer que los individuos están representados por puntos cuya distancia es proporcional a $(1 - S_{ij})^{0,5}$ ”

Mientras que en el caso de Freeman en la sección 16.8.1 re afirma a Gower indicando:

“ Gower, 1966 justificó el uso de PCA en la matriz de covarianza de los datos binarios directamente al mostrar una analogía entre esos resultados y las distancias obtenidas de las frecuencias dispuestas de manera similar a una matriz de Burt donde cada variable estaba representada por una fila y columna.”

Mencionar los trabajos de diferentes instituciones en las cuales se aplica el método de componentes principales sobre un set completo de variables binarias; alguno de ellos son los trabajos de Harttgen & Vollmer, 2011 que emplea variables de activos y McKenzie, 2004 que trabaja una medida de desigualdad empleando también activos del hogar.

Otro aspecto importante a considerar es la conexión entre el método multivariante de análisis de correspondencia múltiple (ACM) y el de componentes principales que para Gower, 1966 son métodos duales, o como se menciona en Murtagh, 2007

“Hay varias formas diferentes de definir y pensar acerca de CA, lo cual es evidente en el redescubrimiento del método por parte de tantos autores diferentes en el siglo pasado. Nos gusta pensar en CA como un tipo de análisis de componentes principales (PCA) de datos categóricos, donde consideramos la definición geométrica de PCA en lugar de su estadística..”

En este punto es importante mencionar que el ACM busca una relación entre las categorías de las variables, mas allá de lo que permite un test Chi-cuadrado, es una

técnica que persigue un enfoque mas visual (técnica gráfica). Por otro lado, el ACP desarrollado en la tesis busca; (1) Identificación de variables relevantes, (2) Identificación de los pesos de estas variables, (3) los cambios que existen de estos grupos de variables en el tiempo.

2.6.1 Un ejemplo con la Encuesta a Hogares 2015, relación entre la pobreza y el indicador basado en CPs

En esta sección se desarrolla un ejemplo del método de componentes principales para datos binarios, se emplea la encuesta a hogares 2015 y se definen 10 variables dicótomas para calcular los componentes. Después se estudia la relación del primer componente principal con variables de bienestar propias de la encuesta; como la pobreza moderada, extrema y el ingreso percapita del hogar.

Las variables seleccionadas para este ejemplo son las de equipamiento del hogar ($act_i = X_i$), estas son capturadas en la sección 8 del cuestionario. Se toman en cuenta a todas los equipamientos y se generan variables dicótomas para cada una de ellas, estas nuevas variables identifican la tenencia. Los equipamientos presentes en la encuesta a hogares 2015 son:

1. Juego de living (act_1)
2. Cocina (act_2)
3. Refrigerador (act_3)
4. Computadora (act_4)
5. Radio (act_5)
6. Equipo de sonido (act_6)
7. Televisor (act_7)
8. Lavadora (act_8)
9. Motocicleta (para uso del hogar) (act_9)
10. Automóvil (para uso del hogar) (act_{10})

El cuadro 2.6 muestra la matriz de varianzas y covarianzas (Σ) para las 10 variables, esta matriz fue calculada considerando los pesos muestrales de la encuesta.

Cuadro 2.6: Matriz de varianzas y covarianzas para las 10 variables de equipamiento

Variables	act_1	act_2	act_3	act_4	act_5	act_6	act_7	act_8	act_9	act_{10}
act_1	0.188									
act_2	0.036	0.132								
act_3	0.078	0.062	0.248							
act_4	0.083	0.028	0.069	0.186						
act_5	-0.013	-0.018	-0.045	-0.015	0.250					
act_6	0.068	0.028	0.062	0.058	-0.049	0.179				
act_7	0.042	0.068	0.077	0.037	-0.025	0.038	0.143			
act_8	0.066	0.020	0.062	0.063	-0.018	0.039	0.020	0.127		
act_9	-0.010	-0.001	0.007	0.001	-0.007	0.006	-0.002	0.003	0.114	
act_{10}	0.042	0.017	0.044	0.046	-0.005	0.031	0.019	0.038	0.003	0.138

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

El cuadro 2.7 presenta los valores propios (λ) y vectores propios (e_i) dada las 10 variables de equipamiento, también, se muestra el PVT para cada uno de los CPs.

Cuadro 2.7: Eigenvalores y eigenvectores de la matriz de covarianza de las variables de equipamiento y el PVT

Variables	Vectores propios									
	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
act_1	0.404	0.246	-0.263	0.226	-0.220	-0.333	0.009	0.580	-0.398	-0.044
act_2	0.239	-0.011	0.424	0.358	-0.094	0.318	-0.003	0.188	0.288	-0.638
act_3	0.509	-0.012	0.488	-0.520	0.110	-0.404	-0.005	-0.194	-0.102	-0.097
act_4	0.378	0.240	-0.337	-0.037	-0.212	0.337	0.511	-0.489	-0.120	-0.108
act_5	-0.229	0.908	0.214	0.068	0.246	-0.070	0.002	-0.057	0.062	0.019
act_6	0.346	-0.101	-0.287	0.425	0.685	-0.185	-0.159	-0.239	0.145	-0.019
act_7	0.283	-0.032	0.448	0.359	-0.056	0.282	0.036	-0.003	-0.112	0.702
act_8	0.281	0.129	-0.210	-0.236	-0.142	-0.027	0.020	0.249	0.804	0.275
act_9	0.010	-0.063	0.001	-0.315	0.577	0.375	0.421	0.479	-0.133	0.009
act_{10}	0.216	0.161	-0.164	-0.283	0.025	0.502	-0.732	0.000	-0.181	-0.021
λ_i	0.547	0.246	0.186	0.135	0.122	0.116	0.108	0.096	0.078	0.067
PVT_i	0.322	0.145	0.109	0.079	0.072	0.068	0.064	0.057	0.046	0.040
Acumulado PVT_i	0.322	0.466	0.576	0.655	0.726	0.794	0.858	0.914	0.960	1.000

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

Se observa que primer componente explica el 32% de la varianza total, esta supera por más del doble la varianza explicada por el segundo componentes, esta varianza es importante dada la cantidad de casos del ejemplo ($N = 10170$). Todos los ponderadores del primer componente son positivos a excepción del que corresponde a la radio (act_5), el act_1 y act_3 son los mas altos, se encuentran entre 0,4 y 0,5, el resto se mueve entre 0,2 y 0,4 excepto act_9 que corresponde a la motocicleta. El primer componente puede ser interpretado como un indicador de riqueza basada en activos, donde, mientras más grande sea el componente más equipamiento tiene el hogar y mientras más pequeño, menos equipamiento tiene el hogar.

El cuadro 2.8 presenta las correlaciones entre los dos primeros componentes y las variables de equipamiento, en este cuadro se confirma la baja relación entre el act_9 y el CP_1 , también, se observa la correlación negativa entre la radio y el CP_1 , es interesante notar que el CP_2 la correlación más alta es con la radio, sin embargo, la alternancia de los

signos en el componente hacen difícil tener una interpretación del componente. En base a la correlación casi nula entre act_9 y el CP_1 , para fines de mejorar el indicador basado en el componente 1 se debe excluir a act_9 y volver a calcular los componentes, esto en base a los criterios observados en 2.5.2.

Cuadro 2.8: Correlaciones entre los 2 primeros componentes y las variables de equipamiento

Variables	CP	
	1	2
act_1	0.690	0.281
act_2	0.486	-0.015
act_3	0.756	-0.011
act_4	0.649	0.277
act_5	-0.340	0.902
act_6	0.605	-0.119
act_7	0.554	-0.041
act_8	0.582	0.180
act_9	0.022	-0.092
act_{10}	0.431	0.215

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

Para finalizar el ejemplo se realizan tres modelos lineales sobre la muestra, esto para ver el comportamiento del R^2 considerando como variable dependiente al ingreso del hogar. Las variables independientes para los modelos son; (1) las diez activos, (2) el primer componente principal y (3) los dos primeros componentes principales. Los cuadros 2.9, 2.10 y 2.11 presentan los resultados respectivamente.

Cuadro 2.9: Modelo 1

	Estimate	Std. Error	t value	$Pr(> t)$
(Intercept)	6.2799	0.0305	205.83	0.0000
act1	0.2192	0.0239	9.18	0.0000
act2	-0.0503	0.0305	-1.65	0.0997
act3	0.2343	0.0209	11.20	0.0000
act4	0.2288	0.0232	9.86	0.0000
act5	-0.0557	0.0181	-3.07	0.0022
act6	0.0908	0.0226	4.01	0.0001
act7	0.4633	0.0307	15.07	0.0000
act8	0.2135	0.0281	7.60	0.0000
act9	0.0876	0.0247	3.55	0.0004
act10	0.0885	0.0250	3.53	0.0004
R2a	0.1696			

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

Cuadro 2.10: Modelo 2

	Estimate	Std. Error	t value	$Pr(> t)$
(Intercept)	6.4097	0.0153	418.28	0.0000
pc1	0.5311	0.0122	43.63	0.0000
R2a	0.1579			

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

Cuadro 2.11: Modelo 3

	Estimate	Std. Error	t value	$Pr(> t)$
(Intercept)	6.3757	0.0176	363.00	0.0000
pc1	0.5290	0.0122	43.45	0.0000
pc2	0.0690	0.0174	3.95	0.0001
R2a	0.1591			

Fuente: Elaboración del autor en base a la Encuesta a Hogares 2015

En el modelo 1 se aprecia que la inclusión de los 10 activos logra un R-cuadrado ajustado (R2a) de casi un 0,17, mientras que para el modelo 2 que incluye únicamente al primer componente logra un **R2a** de 0,16, el tercer modelo es útil para observar la contribución del segundo componente, teniendo una mejora muy pequeña respecto al modelo 2. Tener en cuenta que para el caso del modelo 3, se puede descartar cualquier efecto de colinealidad debido a la independencia entre los 2 componentes, aspecto que no se puede asegurar para el modelo 1. Finalmente, los tres modelos logran un ajuste que ronda entre el 0,16 y el 0,17 y dado que este ejemplo se concentra en los activos del hogar, permite ver como estas características del hogar se relacionan con otras variables de bienestar.

Capítulo 3

Metodología de la investigación

Este capítulo desarrolla la metodología del trabajo, en 3.1 se presenta el diseño de la investigación, en 3.2 se describe los pasos que se seguirán para la construcción del indicador de bienestar dinámico, en 3.4 se detalla los datos que emplearan y los dominios de estudios definidos, finalmente, en 3.5 se presenta el listado de variables que se considerara para la creación del indicador de bienestar planteado.

3.1 Diseño de la investigación

El tipo de investigación que se emplea en el trabajo es de tipo cuantitativa, debido a que se emplean métodos que trabajan a partir de información recolectada por mecanismos basados en censos o encuestas. El área de estudio corresponde al análisis estadístico multivariante, se emplean otros aspectos relacionados al tipo de información tal como ser; inferencia estadística, muestreo, modelos lineales.

En este estudio, se sigue el diseño de investigación¹; **transeccional exploratorio** seguido de un diseño **longitudinal de tendencia**, se pretende en una primera fase comenzar a conocer una variable o un conjunto de variables, una comunidad, un contexto, una situación. Para luego analizar cambios a través del tiempo (en categorías, conceptos, variables o sus relaciones) dentro de alguna población en general.

3.2 Pasos para la definición del “indicador de Bienestar”

Los pasos para la definición de este indicador de Bienestar son:

1. Elegir el grupo de variables que formarán parte del indicador, se recomienda emplear variables que normalmente se capturen de forma sencilla y que su presencia sea habitual en otros estudios similares, la definición de las variables se define en el apartado 3.5.
2. Convertir las variables categóricas a variables binarias, realizar el procedimiento para cada categoría de la variable.

¹?, ?

3. Definir un dominio de interés, esto permitirá comparar el indicador en base a un dominio dado respecto a otros dominios, o con el mismo dominio respecto al tiempo. Esto ayuda ver la evolución del indicador en el tiempo y permite contrastar el estado de bienestar de un dominio a otro.
4. Calcular los componentes principales para la matriz de covarianzas del dominio seleccionado (Componente inicial). En base a lo visto en 2.19.
5. Identificar la cantidad de componentes que explican el 90 % de la varianza en el componente inicial. En base al criterio definido en 2.5.1
6. Eliminar la variable correlacionada con el último componente principal, volver a calcular los componentes principales sin contar con la variable eliminada. Repetir este proceso la misma cantidad de veces que los componentes que quedaron fuera del 90 % en el componente inicial. En base al primer criterio visto en 2.5.2
7. Para los componentes principales resultantes del paso previo (Componente depurado). Calcular las correlaciones entre las variables binarias que permanecen y el primer componente principal, conservar las variables que logran al menos una correlación absoluta del 30 %.
8. Calcular los componentes principales en base a las variables seleccionadas por el punto anterior (indicador de riqueza). En base a lo visto en 2.19.
9. Para medir la evolución del indicador en el tiempo, aplicar los ponderadores del primer componente en el mismo dominio para otros periodos de tiempo y analizar. Siguiendo la relación expresada en 2.19.
10. Para contrastar el indicador con otros dominios, aplicar los ponderadores del primer componente calculado para los otros dominios y analizar. Siguiendo la relación expresada en 2.19.

3.3 CPs como un indicador de monitoreo

La transformación que se logra sobre la matriz de datos X empleando el método de componentes principales, define nuevas variables, cada una de ellas construidas a partir de los eigenvectores, siendo el primer eigen vector el que explica la mayor parte de la variabilidad de la información, así:

$$Y_1 = e_1' \mathbf{X} \quad (3.1)$$

Los valores en e_1 permiten la definición de una nueva variable Y_1 , para el caso de los censos de población, el primer componente principal aplicado a las distintas bases de datos permitirá crear nuevas variables en los censos, tal como:

$$index92_i = e_{1,92}' \mathbf{X}_{92} \quad (3.2)$$

$$index01_i = e_{1,01}' \mathbf{X}_{01} \quad (3.3)$$

$$index12_i = e_{1,12}' \mathbf{X}_{12} \quad (3.4)$$

Donde $index92_i$, $index01_i$, $index12_i$ corresponden a las nuevas variables que son obtenidas por el primer eigenvector (e_1) correspondiente a aplicar CPs a su matrix de datos (X) correspondiente. La información contenida en los distintos eigenvectores puede ser empleada como bases para monitorear los cambios en las nuevas variables según matrices de datos con un soporte común, esto significa en el caso de los censos tener indicadores en base a un año fijo con sus respectivos pesos dados por el primer eigenvector, así, considerando el censo de 1992 como base, se podrá tener indicadores para el censo de 2001 y 2012 con base al censo de 1992, siempre y cuando la matriz de datos tenga un soporte común, las ecuaciones de los indicadores con base en 1992 son:

$$index92_{i,92} = e'_{1,92} \mathbf{X}_{92} \quad (3.5)$$

$$index01_{i,92} = e'_{1,92} \mathbf{X}_{01} \quad (3.6)$$

$$index12_{i,92} = e'_{1,92} \mathbf{X}_{12} \quad (3.7)$$

Si se considera al censo de 2001 como base, se puede tener dos indicadores, estos son:

$$index01_{i,01} = e'_{1,01} \mathbf{X}_{01} \quad (3.8)$$

$$index12_{i,01} = e'_{1,01} \mathbf{X}_{12} \quad (3.9)$$

Para el caso del 2012, dado que no existen censos posteriores se pueden definir indicadores considerando algún dominio (d) de estudio como base respecto a otros para el mismo periodo, por ejemplo, los departamentos. Si calculamos el primer eigenvector para el departamento de La Paz, es posible calcular el indicador para los otros departamentos, teniendo la información de La Paz como base, tal que:

$$index12(d = k)_{i,d=LP} = e'_{1,12,d=LP} \mathbf{X}_{12,d=k} \quad (3.10)$$

Esta forma de emplear el primer eigenvector es una propuesta de este trabajo y esta supone monitorear el indicador de bienestar considerando un año base y compararlo con censos posteriores, empleando las transformaciones dadas por las ecuaciones previas.

3.4 Datos y dominios

Para la aplicación de la metodología se emplea los datos de los tres últimos censos de Población y Vivienda de Bolivia; el CNPV-1992, el CNPV-2001 y el CNPV-2012.

Los dominios de trabajo son a nivel nacional y departamental, para el departamental solo se desarrolla para el CNPV-2012, esto permite realizar los siguientes cruces:

- CNPV 1992 vs CNPV 2001 y CNPV 2012
- CNPV 2001 vs CNPV 2012
- CNPV 2012
- Para el CNPV 2012, a nivel de los 9 departamentos

3.5 Variables

La inclusión de las variables que forman parte del indicador, están vinculadas a la información disponible en los censos y una correspondencia con las dimensiones asociadas al bienestar. El cuadro 3.1 presenta las variables disponibles en los tres censos; 1992, 2001 y 2012. Estas se emplean para la definición del indicador basado en CPs.

Tal como se vio en la figura 2.1, el NBI define 4 necesidades básicas; acceso a vivienda, acceso a servicios sanitarios, acceso a educación y capacidad económica. Este trabajo plantea un indicador de riqueza que incluye variables vinculadas a las siguientes dimensiones:

1. Características de la Vivienda (proxi a: acceso a vivienda)
 - Tipo de vivienda
 - Material de pared
 - Material de piso
 - Material de techo
 - Revoque interior
 - Tenencia de la vivienda
 - Total de cuartos
 - Total de cuartos para dormir
 - Cuarto exclusivo de cocina
2. Características del Hogar
 - Sistema de agua (proxi a: acceso a servicios sanitarios)
 - Procedencia del agua (proxi a: acceso a servicios sanitarios)
 - Servicio sanitario (proxi a: acceso a servicios sanitarios)
 - Uso de servicio sanitario (proxi a: acceso a servicios sanitarios)
 - Tipo de desagüe
 - Fuente de energía eléctrica
 - Combustible de cocina
 - Eliminación de residuos
 - Sexo del jefe de hogar
 - Hogar con niños de 5 años o menos
 - Hogar con adulto de 60 años o más
3. Tenencia de Activos (Proxi a: capacidad económica)
4. Departamento de procedencia

No se incluyen variables vinculadas directamente a la educación, esto debido a dos aspectos:

1. Existen varias alternativas para incluir esta variable; niveles educativos de los miembros o el jefe del hogar, asistencia escolar, rezago escolar, años de educación del jefe o los miembros, etc. Esto depende de una discusión previa respecto el mejor enfoque para incluir esta dimensión y esto se aleja del objetivo central del trabajo. En el caso de las otras dimensiones se incluyeron la mayoría de las variables existentes, principalmente porque existe una correspondencia directa entre las variables y las preguntas empleadas en el censo.
2. La gran parte de las variables incluidas en el cuadro 4.1 son convencionales a la mayoría de los procesos de recolección vinculados a bienestar, sin embargo, existe una diferencia en cuanto a complejidad al momento de incluir variables de educación. Esto motivó la omisión, aunque para futuras aplicaciones del método sugerido, se recomienda tal como se explico en el apartado 3.2 *elegir el grupo de variables que formarán parte del indicador*

Algunas consideraciones a partir del cuadro 4.1:

- Para la variable de número de cuartos, tanto total de cuartos y para dormir se restringe la cota superior a 8 o mas cuartos, esto debido a las opciones de respuesta que se dio en el censo de 1992.
- No se incluye a la variable área dentro de la construcción del indicador, se decide excluir a esta variable con la finalidad de usarla como medida de sensibilidad, tal como se presenta en las figuras 4.10, 4.22 y 4.34. Otro motivo para la exclusión fue la ausencia de un concepto preciso de lo que es área urbana y rural en los censos.

Cuadro 3.1: Variables a nivel de vivienda y hogar en los Censos de Población y Vivienda 1992, 2001 y 2012

Categoría	Grupo	Variables CNPV 1992	Codificación	Variables CNPV 2001	Codificación	Variables CNPV 2012	Codificación
Vivienda	Tipo de Vivienda	Casa independiente	1	Casa, choza, pauchi	11	Casa, choza, pauchi	1
		Departamento	2	Departamento	12	Departamento	2
		Habitación suelta	3	Habitación suelta	13	Habitación suelta	3
		Choza Pahuichi	4	Vivienda improvisada	14	Vivienda improvisada	4
		Local no construido para vivienda	5	Local no construido para vivienda	15	Local no construido para vivienda	5
		Vivienda improvisada	6				
	Material paredes	Adobe revocado	1	Ladrillo, bloques de cemento, hormigón	1	Ladrillo, bloques de cemento, hormigón	1
		Adobe sin revocar o tapial	2	Adobe Tapial	2	Adobe Tapial	2
		Ladrillo, bloques de cemento, hormigón	3	Tabique, quinche	3	Tabique, quinche	3
		Piedra	4	Piedra	4	Piedra	4
Madera		5	Madera	5	Madera	5	
Caña, Palma, troncos		6	Caña, Palma, troncos	6	Caña, Palma, troncos	6	
Otros		7	Otro	7	Otro	7	
Revoque interior de la pared	Tiene	7	Tiene	7	Tiene	7	
	No tiene	8	No tiene	8	No tiene	2	
Material Techos	Calamina o plancha	1	Calamina o plancha	1	Calamina o plancha	1	
	Tejas	2	Tejas	2	Tejas	2	
	Losa de hormigón armado	3	Losa de hormigón armado	3	Losa de hormigón armado	3	
	Paja, caña, palma	4	Paja, caña, palma	4	Paja, caña, palma	4	
	Otros	5	Otros	5	Otros	5	
Material Pisos	Madera	1	Tierra	1	Tierra	1	
	Mosaico o baldosas	2	Tablón de madera	2	Tablón de madera	2	
	Ladrillo	3	Machimbre, parquet	3	Machimbre	3	
	Cemento	4	Alfombra, tapizón	4	Parquet	4	
	Tierra	5	Cemento	5	Cerámica	5	
	Otros	6	Mosaico, baldosa, cerámica	6	Cemento	6	
Tenencia de la vivienda	Propia	1	Ladrillo	7	Mosaico, baldosa	7	
	Alquilada	2	Otra	8	Ladrillo	8	
	Contrato anticretico	3			Otro	9	
	Contrato mixto	4					
	Cedida por servicios	5					
	Cedida por parentesco	6					
	Otra	7					
Características de los cuartos	Total de cuartos		Total de cuartos		Total de cuartos		
	Cuartos para dormir		Cuartos para dormir		Cuartos para dormir		
Cuarto exclusivo de cocina	Tiene	1	Tiene	1	Tiene	1	
	No tiene	2	No tiene	2	No tiene	2	
Servicios	Sistema de agua	Cañería dentro de la vivienda	1	Cañería dentro de la vivienda	6	Cañería dentro de la vivienda	1
		Cañería fuera de la vivienda pero en el lote	2	Cañería fuera de la vivienda pero en el lote	7	Cañería fuera de la vivienda pero en el lote	2
		Cañería fuera del lote	3	No recibe agua por cañería	8	Cañería fuera del lote	3
		No recibe agua por cañería	4			No recibe agua por cañería	4
	Procedencia del agua	Red pública o privada	1	Cañería de red	1	Cañería de red	1
		Pozo o noria	2	Pileta pública	2	Pileta pública	2
		Río, lago, vertiente o acequia	3	Carro repartidor	3	Carro repartidor	3
		Carro repartidor	4	Pozo o noria con bomba	4	Pozo o noria con bomba	4
		Otro	5	Pozo o noria sin bomba	5	Pozo o noria sin bomba	5
				Río, vertiente, acequia	6	Lluvia, río, vertiente, acequia	6
Servicio sanitario	Tiene con descarga instantánea	1	Lago, laguna, curichi	7	Lago, laguna, curichi	7	
	Tiene sin descarga	2	Otra	8	Otra	8	
	No tiene	3					
Uso servicio sanitario	Privado	1	Tiene	1	Tiene y es privado	1	
	Compartido	2	No tiene	2	Tiene y es compartido	2	
Tipo de desagüe	Alcantarillado	1	Privado	3	No tiene	3	
	Cámara séptica	2	Compartido	4			
	Otro	3					
			Alcantarillado	1	Alcantarillado	1	
			Cámara séptica	2	Cámara séptica	2	
Energía eléctrica	Tiene	1	Pozo ciego	3	Pozo ciego	3	
	No tiene	2	Superficie	4	A la calle	4	
					A la quebrada o río	5	
					A lago, laguna, curuchi	6	
Combustible para cocinar	Leña	1	Tiene	5	Red de empresa eléctrica	1	
	Guano, bosta o taquia	2	No tiene	6	Motor propio	2	
	Carbón	3			Panor solar	3	
	Kerosene	4			Otra	4	
	Gas licuado	5			No tiene	5	
	Electricidad	6			Gas domiciliario (por cañería)	1	
	No cocina	7			Gas en garrafá	2	
	Otro	8			Electricidad	3	
Eliminación de residuos					Energía solar	4	
					Leña	5	
					Otro	6	
					Guano, bosta o taquia	6	
					Otro	7	
					No cocina	8	
Activos	Tenencia de activos					Basurero público o contenedor	1
						Carro basurero	2
						Terreno baldío o calle	3
						Río	4
						Quemar	5
						Enterrar	6
						Otra	7
Geográficos	Área	Urbana	1	Urbana	1	Urbana	1
		Rural	2	Rural	2	Rural	2
	Departamento	Chuquisaca	1	Chuquisaca	1	Chuquisaca	1
		La Paz	2	La Paz	2	La Paz	2
		Cochabamba	3	Cochabamba	3	Cochabamba	3
		Oruro	4	Oruro	4	Oruro	4
		Potosí	5	Potosí	5	Potosí	5
		Tarija	6	Tarija	6	Tarija	6
		Santa Cruz	7	Santa Cruz	7	Santa Cruz	7
		Beni	8	Beni	8	Beni	8
Pando	9	Pando	9	Pando	9		
Hogar	Característica del hogar	Jefe del hogar hombre		Jefe del hogar hombre		Jefe del hogar hombre	
		Jefe del hogar mujer		Jefe del hogar mujer		Jefe del hogar mujer	
		Hogar con niños en primera infancia (5 años o menos)		Hogar con niños en primera infancia (5 años o menos)		Hogar con niños en primera infancia (5 años o menos)	
		Hogar con adulto mayor (60 años o más)		Hogar con adulto mayor (60 años o más)		Hogar con adulto mayor (60 años o más)	

Capítulo 4

Resultados, conclusiones y recomendaciones

En este capítulo se presentan los resultados de la metodología planteada en base a los datos y dominios descritos en 3.4, también, se incluyen las conclusiones y recomendaciones.

El cuadro 4.1 presenta las variables y ponderadores resultantes para cada periodo base de los censos.

Cuadro 4.1: Variables y ponderadores para el indicador de riqueza según el censo base

Nro.	Base 1992			Base 2001			Base 2012		
	Variable	Descripción	Valor	Variable	Descripción	Valor	Variable	Descripción	Valor
1	mps1	Piso de tierra	-0.36	mps1	Piso de tierra	-0.36	mps1	Piso de tierra	-0.3
2	ec1	Combustible para cocinar: Leña	-0.35	ted1	Un cuarto para dormir	-0.18	ec5	Combustible para cocinar: Leña	-0.27
3	mt4	Techo de paja, caña o palma	-0.26	mp2	Pared de adobe	-0.17	er5	Eliminación de residuos: quemar	-0.18
4	tnv1	Tenencia de vivienda: propia	-0.2	tnv1	Tenencia de vivienda: propia	-0.14	mp2	Pared de adobe tapial	-0.17
5	pa3	Procedencia de agua: Pozo o noria	-0.16	pa5	Procedencia de agua: pozo o noria sin bomba	-0.11	tnv1	Tenencia de vivienda: propia	-0.14
6	tv1	Casa Independiente	-0.15	mps6	Piso de mosaico, baldosa, cerámica	0.08	pa6	Procedencia de agua: lluvia, río, vertiente	-0.11
7	ss2	Con servicio sanitario compartido	0.11	act4	Vivienda con vehículo	0.11	tnv2	Tenencia de vivienda: alquilada	0.09
8	tnv2	Tenencia de vivienda: Alquilada	0.13	mt2	Techo de tejas	0.13	act4	Vivienda con servicio de internet	0.09
9	sa2	Cañería fuera de la vivienda pero en el lote	0.14	act1	Vivienda con radio, equipo de sonido	0.15	mt2	Techo de tejas	0.11
10	ss1	Con servicio sanitario privado	0.14	rev1	Con revoque interior de paredes	0.18	act6	Vivienda con vehículo	0.11
11	mp2	Pared de ladrillo, cemento	0.17	mps4	Piso de Cemento	0.19	mps5	Piso de cerámica	0.12
12	mt1	Techo de calamina o plancha	0.18	ss1	Con servicio sanitario privado	0.21	rev1	Con revoque interior de paredes	0.17
13	mps3	Piso de cemento	0.19	act5	Vivienda con teléfono o celular	0.23	ss1	Con servicio sanitario privado	0.17
14	td1	Desagüe con alcantarillado	0.2	td1	Desagüe con alcantarillado	0.26	act3	Vivienda con computadora	0.19
15	pa1	Procedencia de agua: red pública/privada	0.35	pa1	Procedencia de agua: Cañería de red	0.3	ec2	Combustible para cocinar: Gas en garrafa	0.24
16	ee1	Con energía eléctrica	0.37	ee1	Con energía eléctrica	0.36	pa1	Procedencia de agua: cañería de red	0.25
17	ec4	Combustible para cocinar: Gas licuado	0.37	ec4	Combustible para cocinar: Gas (garrafa/cañería)	0.36	ee1	Con energía eléctrica	0.25
18				act2	Vivienda con televisor	0.37	er2	Eliminación de residuos: carro basurero	0.27
19							sa1	Sistema de agua: cañería dentro de la vivienda	0.28
20							td1	Desagüe con alcantarillado	0.29
21							act5	Vivienda con servicio de telefonía fija o celular	0.29
22							act2	Vivienda con televisor	0.31

4.1 Resultados

Los resultados se presentan según los dominios y periodos descritos; se inicia con la serie completa de los censos 1992 a 2012, teniendo como base al censo de 1992, luego se presenta los resultados para el periodo 2001 a 2012, definiendo como base el censo de 2001, finalmente se presenta los resultados para el censo de 2012.

Los resultados se presentan según los pasos (ver 3.2) descritos en el capítulo de la metodología, al finalizar se incluye la comparación por periodo según el periodo base, esto se realiza a nivel global y a nivel municipal.

4.1.1 CNPV 1992 vs CNPV 2001 y CNPV 2012

Se inicia con las variables del censo de 1992 descritas en el cuadro 3.1, cada una de ellas son llevadas a variables binarias. Se inicia con 63 variables, la figura 4.1 presenta un gráfico de correlaciones de estas 63 variables iniciales.

A partir de estas 63 variables se calcula el componente principal inicial, la figura 4.2 presenta el Scree plot de los eigen valores, también, se presenta en la figura 4.3 el porcentaje acumulado de la varianza.

Figura 4.1: Correlaciones entre las primeras 63 variables del censo de 1992

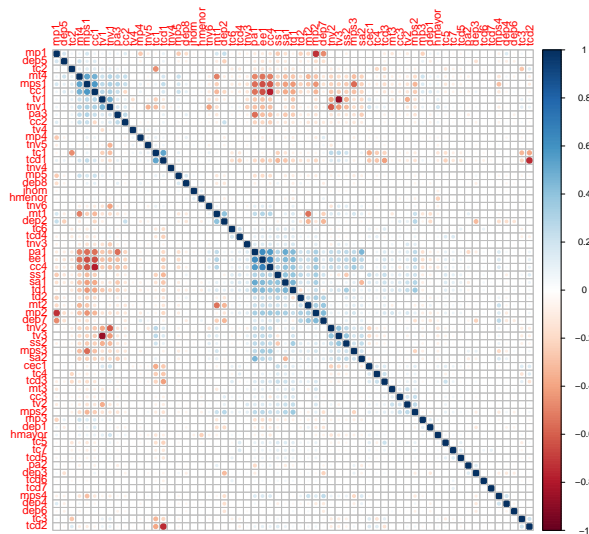


Figura 4.2: Scree Plot del componente principal inicial para el censo de 1992

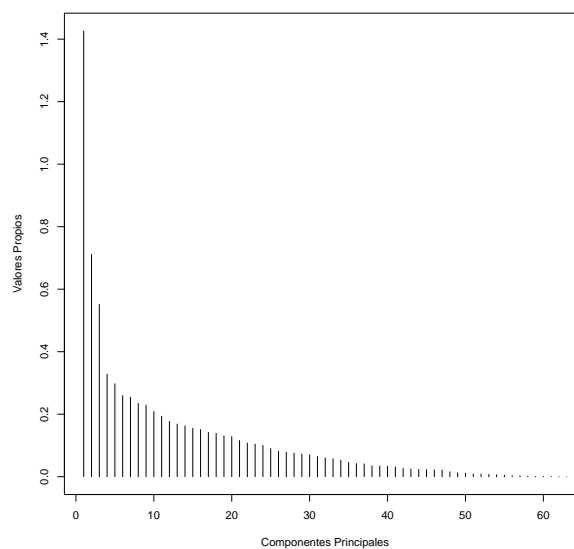
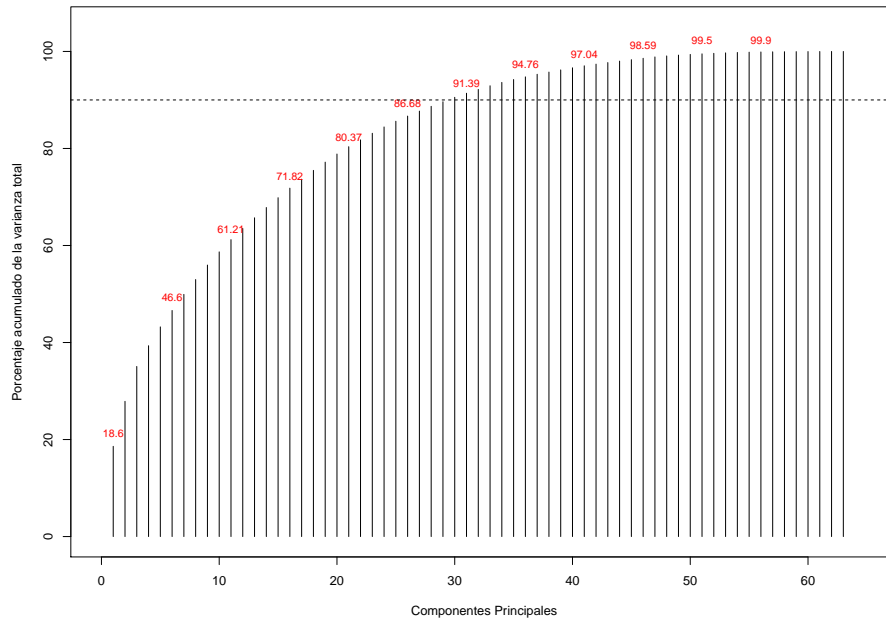


Figura 4.3: Porcentaje acumulada de la varianza para el componente inicial para el censo de 1992

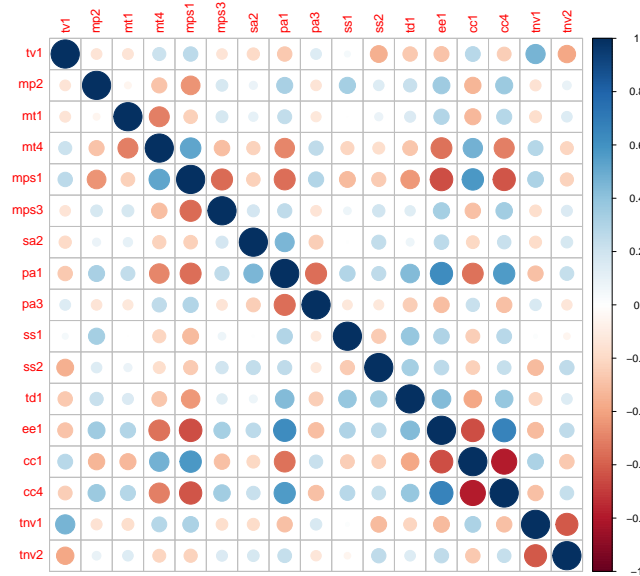


Se aprecia que la gran cantidad de las variables conduce a que la varianza se encuentre dispersa, por ello es necesario aplicar la metodología para reducir las variables que no estén aportando en términos de la varianza. Para ello se sigue los siguientes tres criterios:

1. **En base al componente inicial se identifica el número de componentes principales que explican el 90 % de la varianza.** Se identificaron 29 componentes, es decir, los 34 componentes finales explican el 10 de la varianza.
2. **Siguiendo lo descrito en 2.5.2 se descartan 34 variables asociadas con el último componente principal,** para ello se sigue un proceso iterativo de cálculo del componente a partir de las variables relevantes y el descarte de las variables asociadas con el último componente principal.
3. Se seleccionan las variables que logren al menos una correlación absoluta superior a 30 % con el primer componente principal, del componente principal basado en las 29 variables restantes del criterio anterior (componente depurado). Finalmente, se calcula el componente principal en base a las variables que cumplen el criterio de la correlación (indicador de riqueza).

Aplicando los criterios, se terminó seleccionando un total de 17 variables. La figura 4.4 presenta las correlaciones entre las variables finales para el indicador de riqueza.

Figura 4.4: Correlaciones entre las 17 variables resultantes de los criterios, para el censo de 1992



La figura 4.5 presenta el screeplot para el componente principal final, mientras que en 4.6 se presenta el porcentaje acumulado de la varianza para el componente final.

Figura 4.5: Scree Plot del componente principal inicial para el censo de 1992

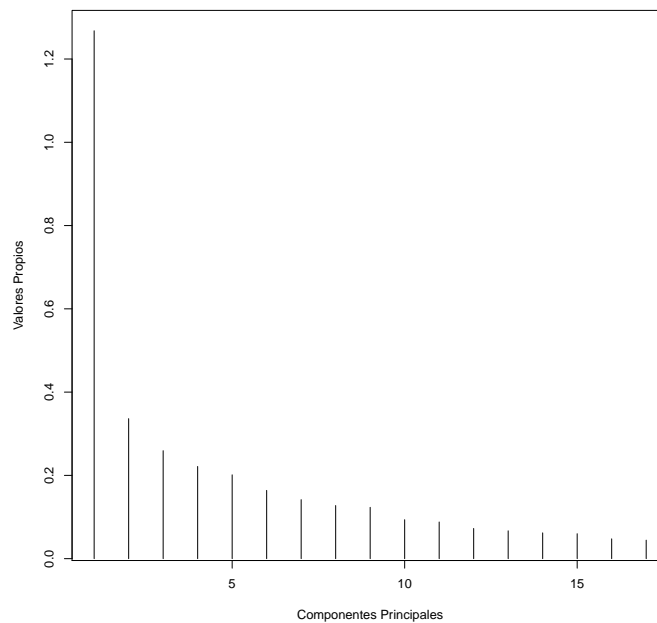
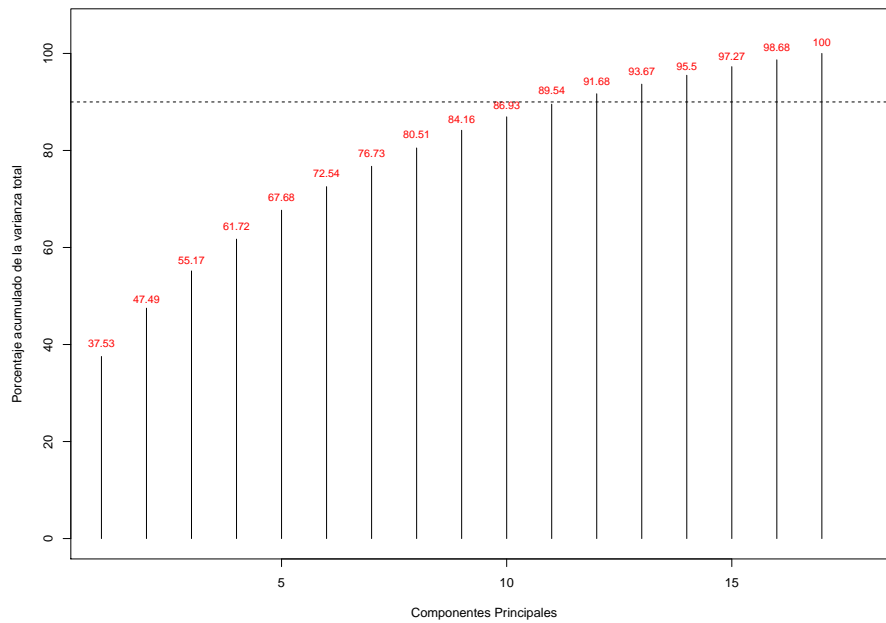
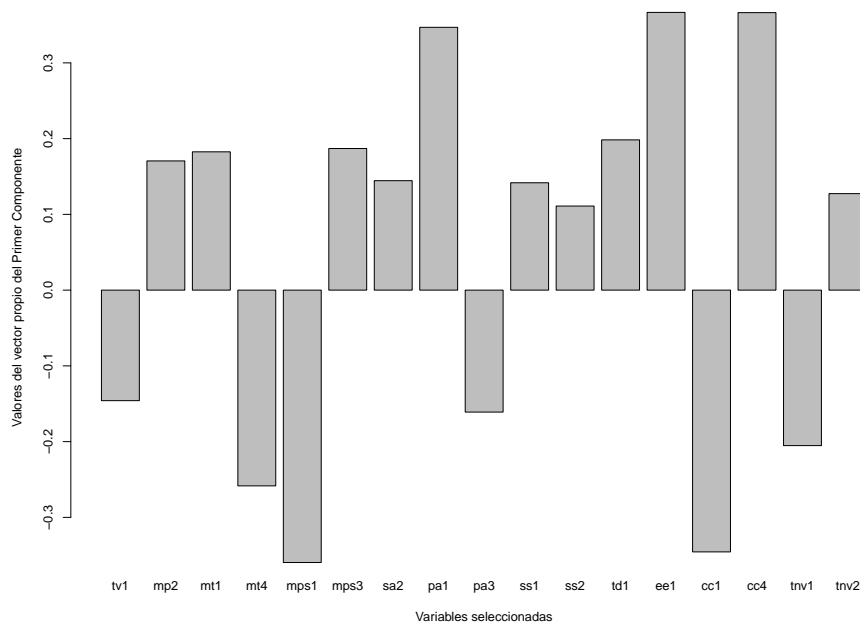


Figura 4.6: Porcentaje acumulada de la varianza para el componente inicial para el censo de 1992



La figura 4.7 presenta la dirección de las variables asociadas al primer componente según las variables finales.

Figura 4.7: Eigenvalores y dirección para el primer componente para el censo de 1992



Los cuadros 4.2 y 4.3 presenta los eigenvalores y eigenvectores respectivamente, se indica la variable asociada a cada valor dentro de los eigenvectores.

Cuadro 4.2: Eigenvalores y % acumulado para el indicador de riqueza, 1992

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}
Eigenvalores	1.27	0.34	0.26	0.22	0.20	0.16	0.14	0.13	0.12	0.09	0.09	0.07	0.07	0.06	0.06	0.05	0.04
% acumulado	37.53	47.49	55.17	61.72	67.68	72.54	76.73	80.51	84.16	86.93	89.54	91.69	93.67	95.50	97.27	98.68	100.00

Cuadro 4.3: Eigenvectores para el indicador de riqueza, 1992

	tv1	mp2	mt1	mt4	mps1	mps3	sa2	pa1	pa3	ss1	ss2	td1	ee1	cc1	cc4	tnv1	tnv2
eigenvector 1	-0.15	0.17	0.18	-0.26	-0.36	0.19	0.14	0.35	-0.16	0.14	0.11	0.20	0.37	-0.35	0.37	-0.20	0.13
eigenvector 2	-0.33	-0.22	0.22	-0.01	0.13	0.01	0.15	-0.06	0.01	-0.52	0.26	-0.11	-0.08	0.05	-0.11	-0.50	0.36
eigenvector 3	-0.20	0.28	-0.77	0.29	-0.04	-0.08	0.10	0.11	-0.09	0.11	0.14	0.13	-0.02	0.09	-0.06	-0.27	0.17
eigenvector 4	0.09	-0.18	0.04	0.00	0.18	-0.19	0.53	0.49	-0.50	-0.07	0.02	-0.01	-0.02	0.18	-0.13	0.22	-0.12
eigenvector 5	-0.12	-0.04	0.21	-0.04	0.32	-0.68	-0.25	0.06	-0.04	0.35	-0.13	0.29	0.03	-0.08	-0.01	-0.23	0.13
eigenvector 6	0.05	-0.03	-0.20	0.18	0.22	-0.34	0.05	-0.01	0.15	-0.41	0.16	-0.12	0.14	-0.49	0.45	0.23	-0.12
eigenvector 7	0.22	0.24	-0.04	-0.15	0.10	-0.10	0.28	0.02	0.09	0.15	-0.43	-0.65	-0.02	-0.08	0.10	-0.26	0.21
eigenvector 8	0.20	-0.30	-0.12	0.18	-0.02	0.17	-0.51	0.08	-0.58	-0.11	-0.23	-0.12	-0.07	-0.10	0.19	-0.15	0.14
eigenvector 9	-0.05	0.65	0.08	-0.30	-0.04	-0.21	-0.34	0.03	-0.30	-0.34	0.16	-0.16	-0.04	0.16	-0.09	0.09	-0.11
eigenvector 10	0.63	-0.16	-0.17	-0.22	-0.07	-0.14	-0.08	0.05	0.16	-0.15	0.18	0.09	0.51	0.23	-0.16	-0.16	0.13
eigenvector 11	0.56	0.25	0.18	0.12	0.02	0.04	0.15	-0.01	-0.00	0.01	0.25	0.30	-0.57	-0.13	0.13	-0.17	0.11
eigenvector 12	-0.07	-0.32	-0.34	-0.67	-0.14	-0.14	0.05	-0.15	-0.04	0.02	0.03	0.05	-0.38	0.03	0.24	-0.10	-0.21
eigenvector 13	0.02	-0.04	0.13	0.33	-0.49	-0.27	0.05	0.04	0.05	-0.13	-0.15	-0.06	-0.01	0.05	-0.03	-0.41	-0.59
eigenvector 14	0.04	0.10	0.04	0.01	0.43	0.20	0.15	-0.46	-0.31	0.18	0.18	0.00	0.29	0.10	0.19	-0.28	-0.38
eigenvector 15	0.00	-0.01	0.08	0.12	-0.42	-0.30	0.25	-0.57	-0.30	-0.00	-0.06	0.06	0.09	0.11	0.12	0.24	0.36
eigenvector 16	-0.06	-0.06	0.11	0.16	-0.02	-0.05	-0.12	0.21	0.17	0.15	0.26	-0.21	-0.05	0.61	0.59	0.05	0.08
eigenvector 17	-0.02	0.17	-0.03	-0.04	0.16	0.11	0.10	0.02	0.10	-0.40	-0.61	0.47	0.05	0.28	0.28	-0.03	0.01

A partir de los valores dados por el primer componente resultante de estas 17 variables, se calcula el Indicador de riqueza para los tres censos, la figura 4.8 presenta los histogramas y la densidad del indicador de riqueza, en las figuras 4.9, 4.10 y 4.11 presentan los diagramas de caja para el indicador de riqueza global, por área y departamento para cada censo respectivamente.

Figura 4.8: Densidad e histograma a partir del indicador para 1992

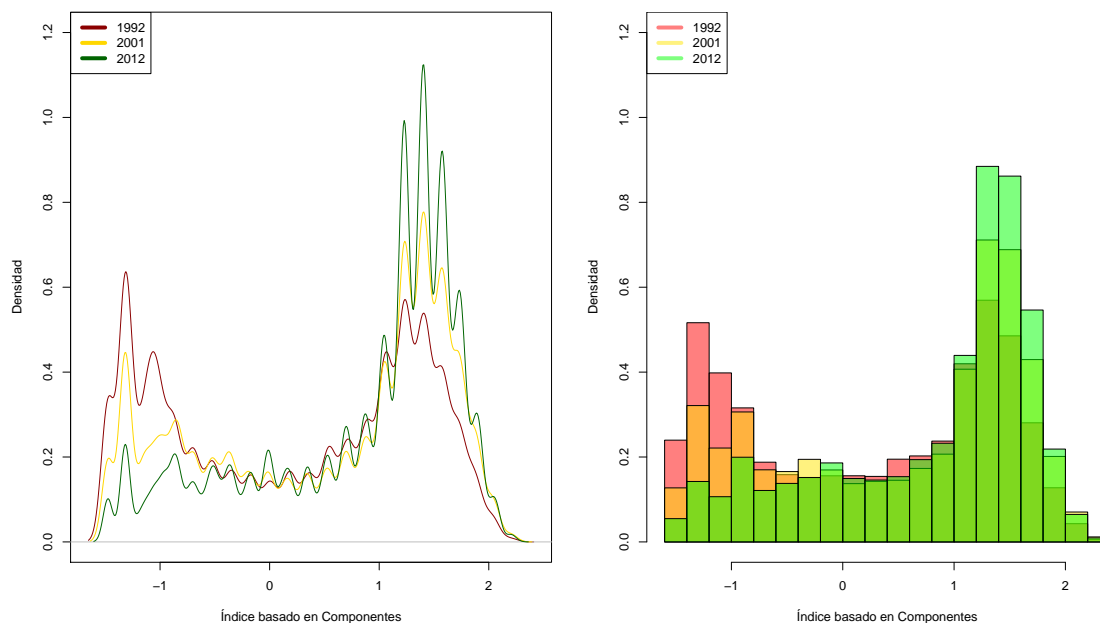


Figura 4.9: Diagramas de caja por Censo del indicador de riqueza

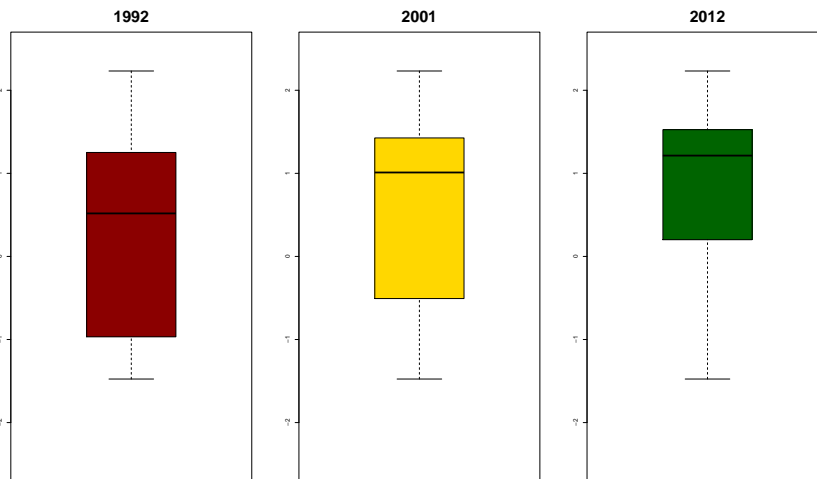


Figura 4.10: Diagrama de caja por Censo y área del indicador de riqueza

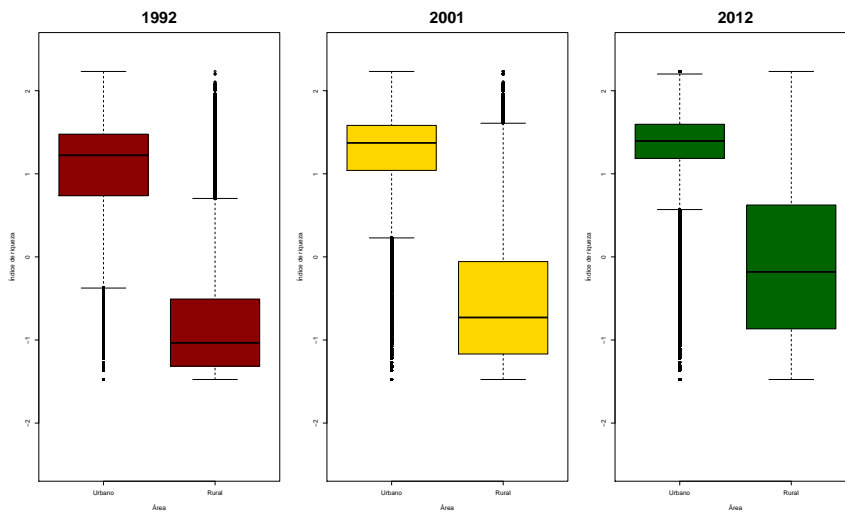
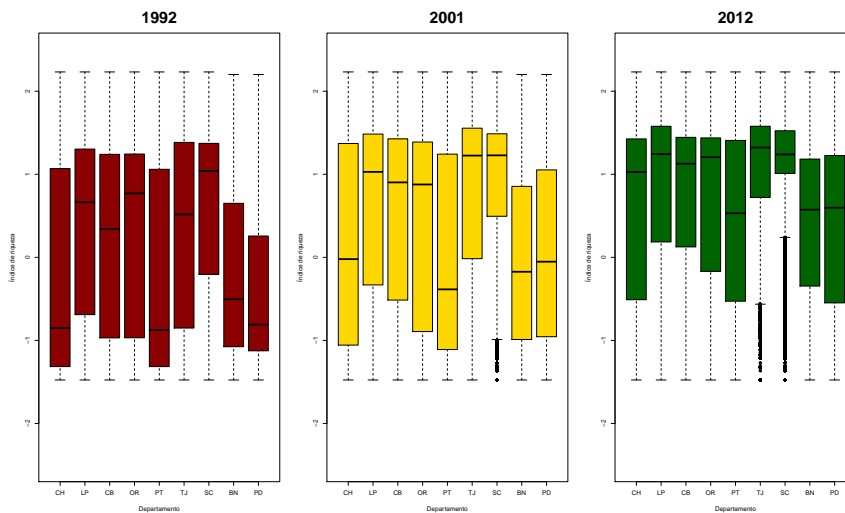
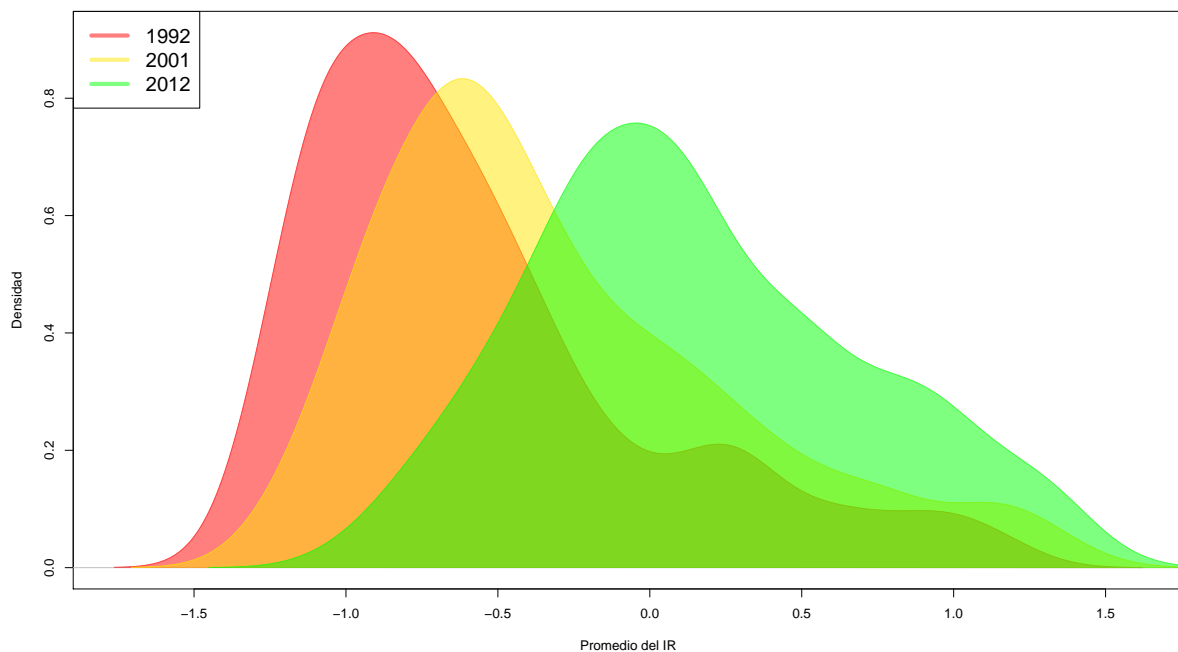


Figura 4.11: Diagrama de caja por Censo y departamento del indicador de riqueza



Estas figuras permiten visualizar los cambios en riqueza a partir de la definición dada en el censo de 1992, de manera global para los censo 2001 y 2012 las viviendas tienden a concentrarse en valores altos del indicador, a nivel del área lo urbano se empieza a concentrar según avanza el tiempo, mientras que para el caso rural existe un desplazamiento hacia valores altos del indicador sin dejar la cola y sin alcanzar los valores del área urbana. A nivel de los departamentos, todos empiezan a concentrarse en valores altos a medida que pasa el tiempo, siendo Tarija y Santa Cruz los departamentos con una mayor variación. A partir del indicador de riqueza se obtiene el promedio del indicador a nivel municipal, la figura 4.12 presenta la densidad del indicador a nivel municipal, en este se aprecia claramente el desplazamiento de la riqueza a medida que pasa el tiempo, esto dada la definición a partir del censo de 1992.

Figura 4.12: Evolución del indicador de riqueza a nivel municipal en base al censo de 1992



4.1.2 CNPV 2001 vs CNPV 2012

Se inicia con las variables del censo de 2001 descritas en el cuadro 3.1, cada una de ellas son llevadas a variables binarias. Se inicia con 77 variables, la figura 4.13 presenta un gráfico de correlaciones de estas 77 variables iniciales.

A partir de estas 77 variables se calcula el componente principal inicial, la figura 4.14 presenta el Scree plot de los eigen valores, también, se presenta en la figura 4.15 el porcentaje acumulado de la varianza.

Figura 4.13: Correlaciones entre las primeras 77 variables del censo de 2001

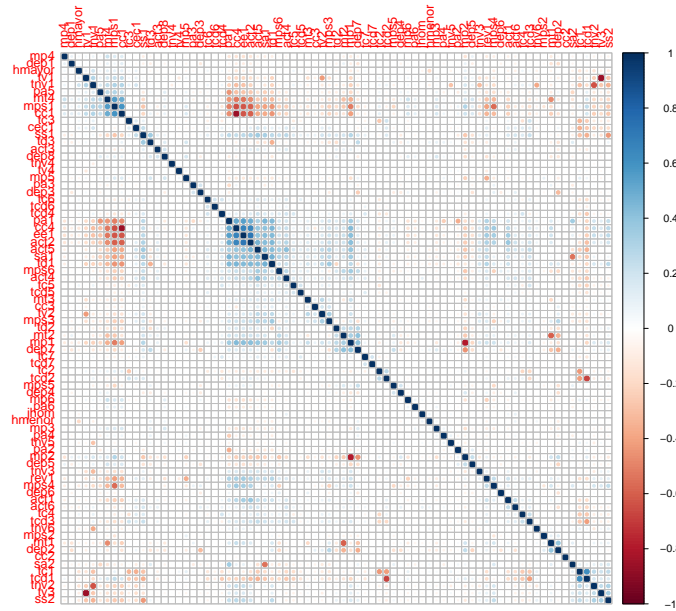


Figura 4.14: Scree Plot del componente principal inicial para el censo de 2001

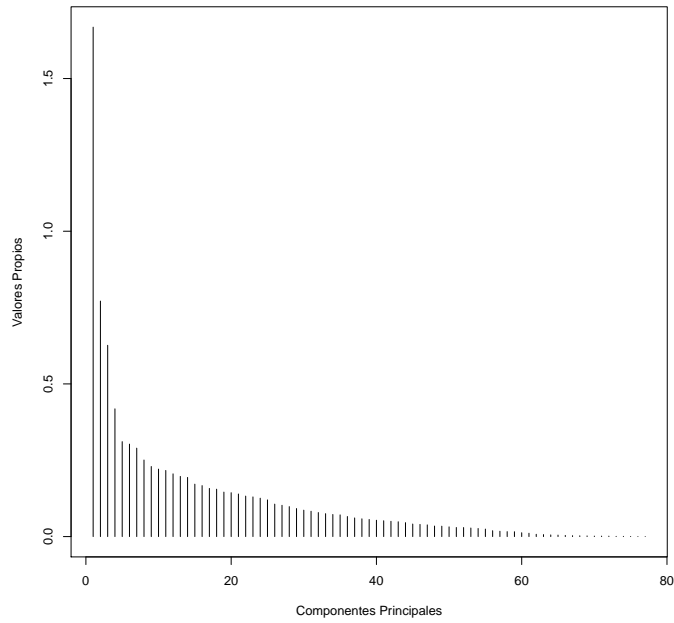
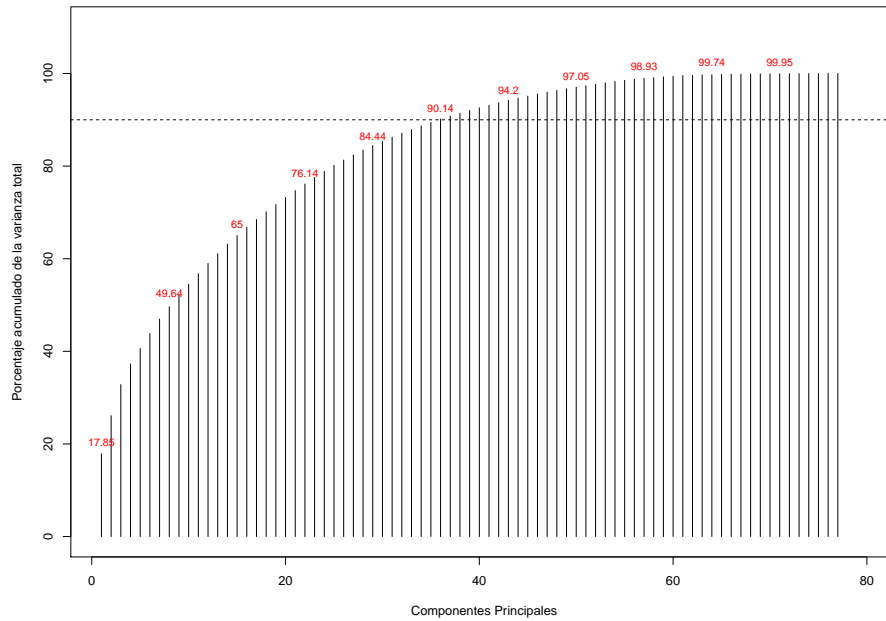


Figura 4.15: Porcentaje acumulada de la varianza para el componente inicial para el censo de 2001

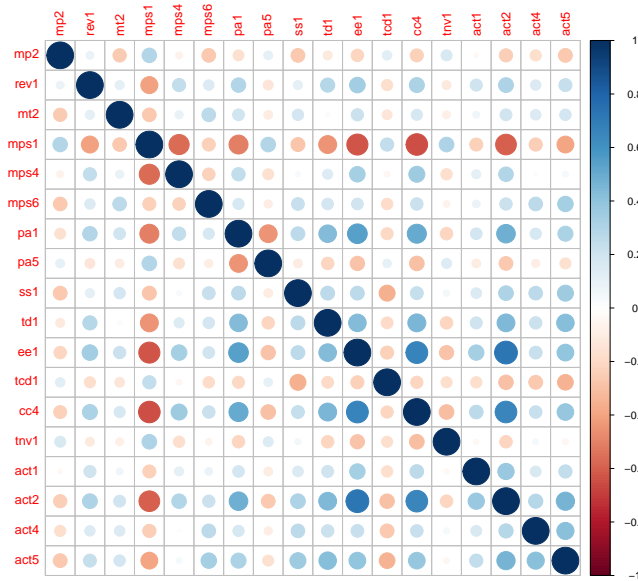


Se aprecia que la gran cantidad de las variables conduce a que la varianza se encuentre dispersa, por ello es necesario aplicar la metodología para reducir las variables que no estén aportando en términos de la varianza. Para ello se sigue los siguientes tres criterios:

1. **En base al componente inicial se identifica el número de componentes principales que explican el 90 % de la varianza.** Se identificaron 29 componentes, es decir, los 34 componentes finales explican el 10 de la varianza.
2. **Siguiendo lo descrito en 2.5.2 se descartan 34 variables asociadas con el último componente principal,** para ello se sigue un proceso iterativo de cálculo del componente a partir de las variables relevantes y el descarte de las variables asociadas con el último componente principal.
3. Se seleccionan las variables que logren al menos una correlación absoluta superior a 30 % con el primer componente principal, del componente principal basado en las 29 variables restantes del criterio anterior (componente depurado). Finalmente, se calcula el componente principal en base a las variables que cumplen el criterio de la correlación (índice de riqueza).

Aplicando los criterios, se terminó seleccionando un total de 18 variables. La figura 4.16 presenta las correlaciones entre las variables finales para el indicador de riqueza.

Figura 4.16: Correlaciones entre las 18 variables resultantes de los criterios, para el censo de 2001



La figura 4.17 presenta el screeplot para el componente principal final, mientras que en 4.18 se presenta el porcentaje acumulado de la varianza para el componente final.

Figura 4.17: Scree Plot del componente principal inicial para el censo de 2001

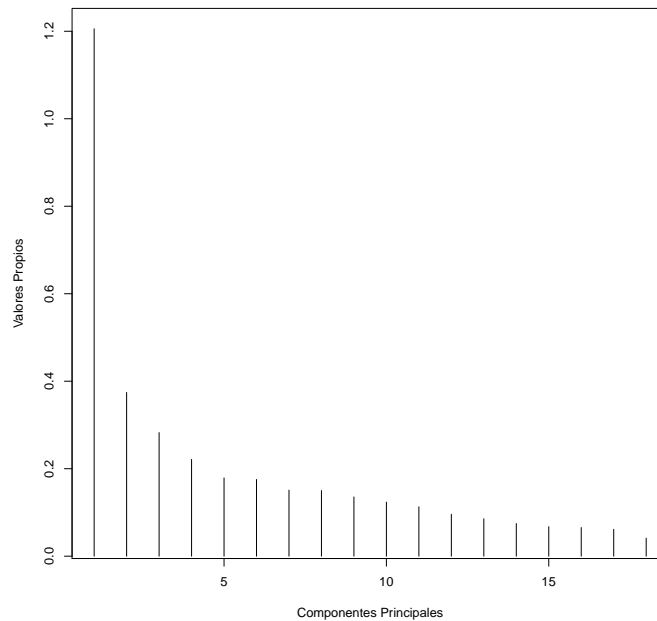
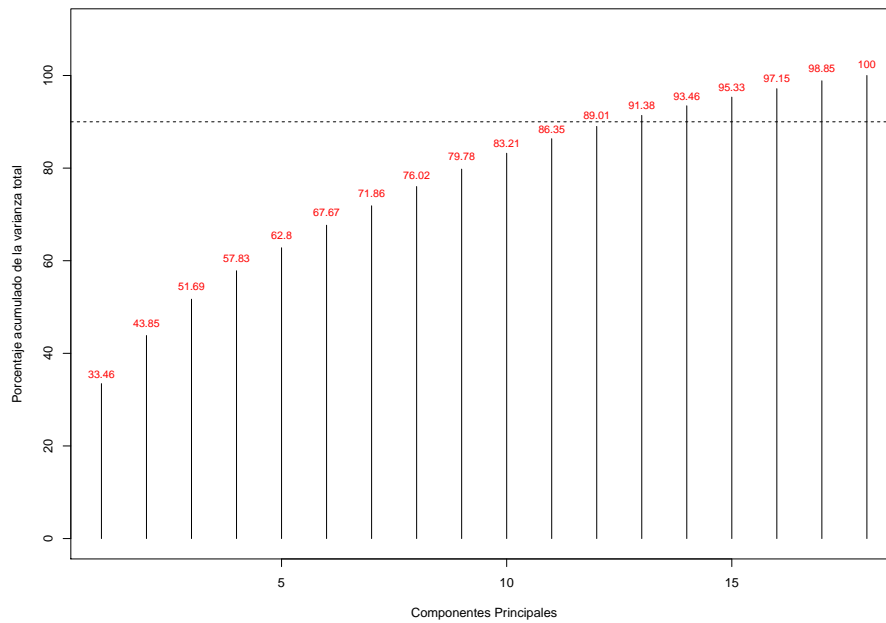
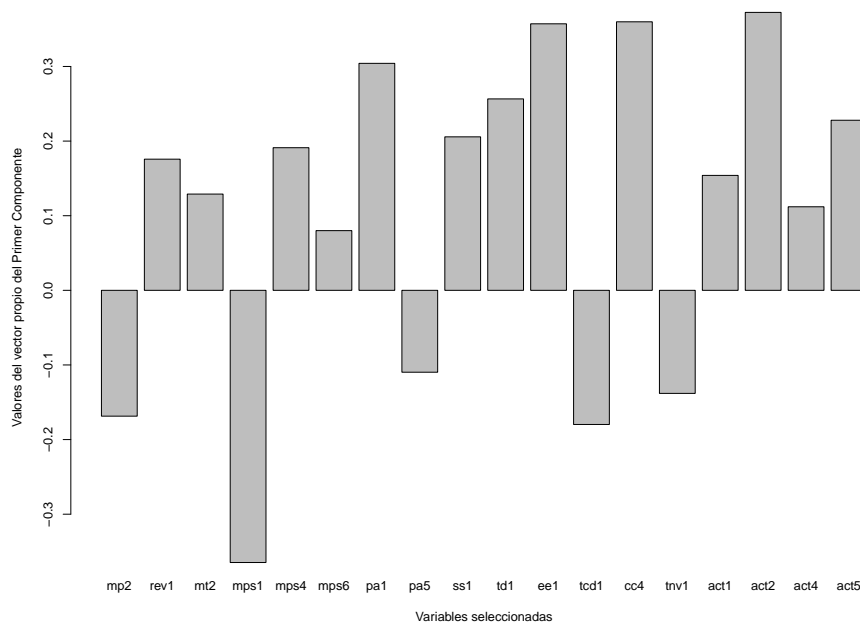


Figura 4.18: Porcentaje acumulada de la varianza para el componente inicial para el censo de 2001



La figura 4.19 presenta la dirección de las variables asociadas al primer componente según las variables finales.

Figura 4.19: Eigenvalores y dirección para el primer componente para el censo de 2001



Los cuadros 4.4 y 4.5 presenta los eigenvalores y eigenvectores respectivamente, se indica la variable asociada a cada valor dentro de los eigenvectores.

Cuadro 4.4: Eigenvalores y % acumulado para el indicador de riqueza, 2001

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}
Eigenvalores	1.21	0.37	0.28	0.22	0.18	0.18	0.15	0.15	0.14	0.12	0.11	0.10	0.09	0.07	0.07	0.07	0.06	0.04
% acumulado	33.46	43.85	51.69	57.83	62.80	67.67	71.86	76.02	79.78	83.21	86.34	89.01	91.38	93.46	95.33	97.15	98.86	100.00

Cuadro 4.5: Eigenvectores para el indicador de riqueza, 2001

	mp2	rev1	mt2	mps1	mps4	mps6	pa1	pa5	ss1	td1	ee1	tcd1	cc4	tnv1	act1	act2	act4	act5
eigenvector 1	-0.17	0.18	0.13	-0.36	0.19	0.08	0.30	-0.11	0.21	0.26	0.36	-0.18	0.36	-0.14	0.15	0.37	0.11	0.23
eigenvector 2	0.21	0.09	-0.12	-0.16	0.37	-0.15	0.07	-0.05	-0.47	0.00	0.11	0.47	0.15	-0.41	-0.05	0.01	-0.19	-0.25
eigenvector 3	0.71	0.31	-0.32	0.04	0.07	-0.10	0.04	0.00	-0.06	0.14	0.05	-0.26	-0.00	0.35	0.23	0.07	0.02	0.01
eigenvector 4	-0.08	0.07	0.31	-0.25	0.64	-0.13	-0.14	0.04	0.04	-0.41	-0.06	-0.26	-0.06	0.32	-0.07	-0.08	-0.02	-0.18
eigenvector 5	0.18	0.25	0.67	0.09	-0.30	0.17	-0.04	0.02	-0.34	-0.27	0.10	0.05	-0.03	-0.05	0.32	0.10	0.07	0.04
eigenvector 6	-0.26	-0.28	-0.32	0.05	0.06	-0.08	-0.45	0.16	-0.20	-0.17	0.15	0.00	0.12	0.07	0.55	0.31	0.03	0.05
eigenvector 7	0.03	-0.45	0.03	0.19	-0.09	-0.10	0.56	-0.27	0.00	-0.26	0.21	0.07	0.10	0.32	0.06	0.16	-0.15	-0.27
eigenvector 8	0.15	0.07	0.11	0.03	0.11	-0.07	-0.04	0.06	0.69	-0.03	-0.01	0.48	-0.15	-0.08	0.41	-0.04	-0.07	-0.12
eigenvector 9	0.14	-0.10	-0.03	0.10	-0.07	-0.07	-0.01	0.00	0.17	-0.18	0.05	-0.59	-0.05	-0.63	0.10	-0.01	-0.21	-0.29
eigenvector 10	0.32	-0.13	0.09	0.04	-0.12	0.03	-0.37	0.23	0.23	-0.16	0.28	0.11	0.33	0.04	-0.52	0.33	-0.01	-0.04
eigenvector 11	0.30	-0.63	0.38	-0.02	0.22	-0.12	-0.09	0.03	-0.09	0.46	-0.11	-0.06	-0.08	-0.06	0.08	-0.02	0.07	0.19
eigenvector 12	-0.21	0.14	0.19	0.01	-0.13	-0.00	-0.14	0.18	-0.05	0.51	0.07	-0.07	0.09	0.25	0.05	0.02	-0.51	-0.47
eigenvector 13	-0.10	0.11	0.00	0.17	0.10	-0.20	0.07	0.12	-0.05	0.07	0.34	0.01	-0.73	-0.04	-0.18	0.41	-0.04	0.10
eigenvector 14	-0.03	-0.03	-0.02	-0.05	-0.03	0.07	0.00	0.07	-0.02	0.16	0.11	-0.01	-0.08	-0.01	0.00	-0.01	0.75	-0.61
eigenvector 15	-0.04	-0.01	0.01	0.15	0.06	-0.12	0.38	0.82	-0.04	-0.06	0.11	-0.02	0.16	-0.02	0.04	-0.26	0.05	0.12
eigenvector 16	0.14	-0.23	-0.14	-0.65	-0.17	0.45	0.04	0.17	-0.03	-0.10	0.23	0.02	-0.32	0.06	0.04	-0.16	-0.18	-0.02
eigenvector 17	-0.04	0.02	0.02	0.17	0.04	-0.12	-0.21	-0.26	0.01	0.05	0.69	-0.01	0.03	0.03	0.01	-0.58	0.03	0.12
eigenvector 18	0.02	-0.02	-0.06	0.47	0.41	0.77	0.02	0.00	0.00	0.06	0.04	-0.01	-0.00	-0.02	0.00	0.04	-0.06	-0.02

A partir de los valores dados por el primer componente resultante de estas 17 variables, se calcula el indicador de riqueza para los tres censos, la figura 4.20 presenta los histogramas y la densidad del indicador de riqueza, en las figuras 4.21, 4.22 y 4.23 presentan los diagramas de caja para el indicador de riqueza global, por área y departamento para cada censo respectivamente.

Figura 4.20: Densidad e histograma a partir del indicador para 2001

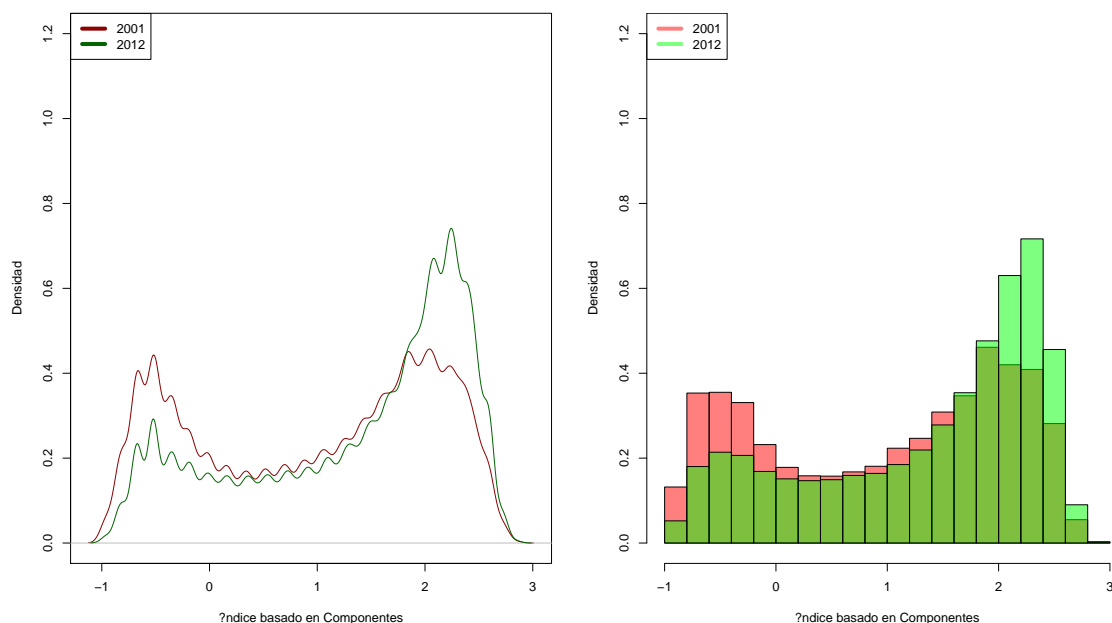


Figura 4.21: Diagrama de caja por Censo del indicador de riqueza

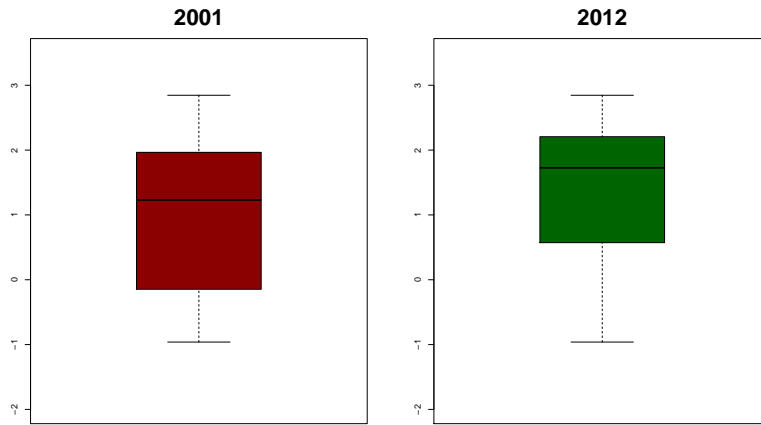


Figura 4.22: Diagrama de caja por Censo y área del indicador de riqueza

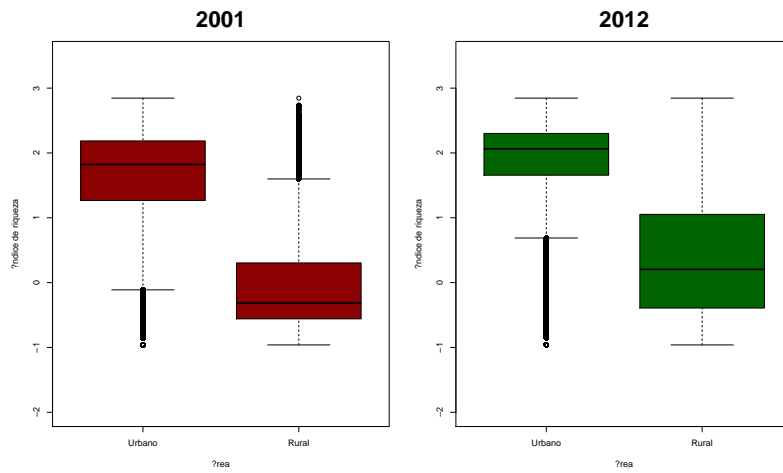
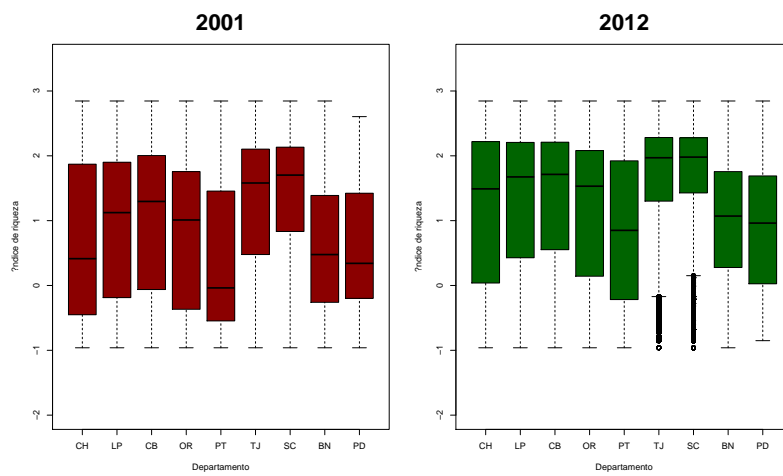
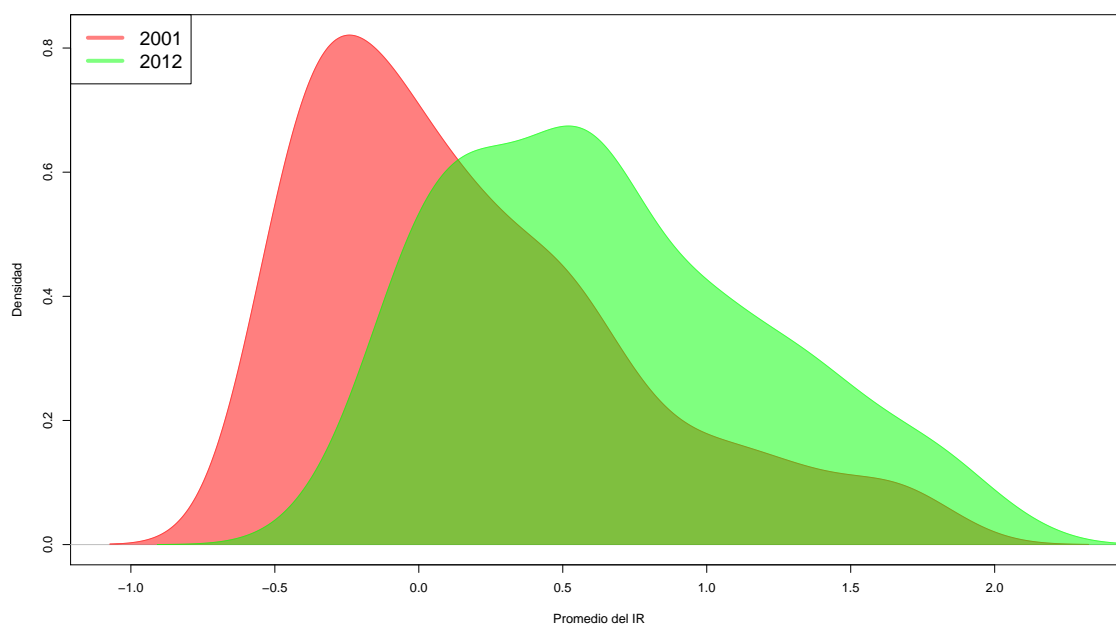


Figura 4.23: Diagrama de caja por Censo y departamento del indicador de riqueza



Estas figuras permiten visualizar los cambios en riqueza a partir de la definición dada en el censo de 2001, de manera global para el censo de 2012 las viviendas tienden levemente a concentrarse en valores altos del indicador, a nivel del área lo urbano se empieza a concentrar según avanza el tiempo, mientras que para el caso rural existe un desplazamiento hacia valores altos del indicador sin dejar la cola y sin alcanzar los valores del área urbana. A nivel de los departamentos, todos empiezan a concentrarse en valores altos del indicador, sin embargo, lo hacen de manera mas lenta que los visto con el censo de 1992. Nuevamente Tarija y Santa Cruz son los departamentos con una mayor variación. A partir del indicador de riqueza se obtiene el promedio del indicador a nivel municipal, la figura 4.24 presenta la densidad del indicador a nivel municipal, en este se aprecia claramente el desplazamiento de la riqueza a medida que pasa el tiempo, esto dada la definición a partir del censo de 2001.

Figura 4.24: Evolución del indicador de riqueza a nivel municipal en base al censo de 2001

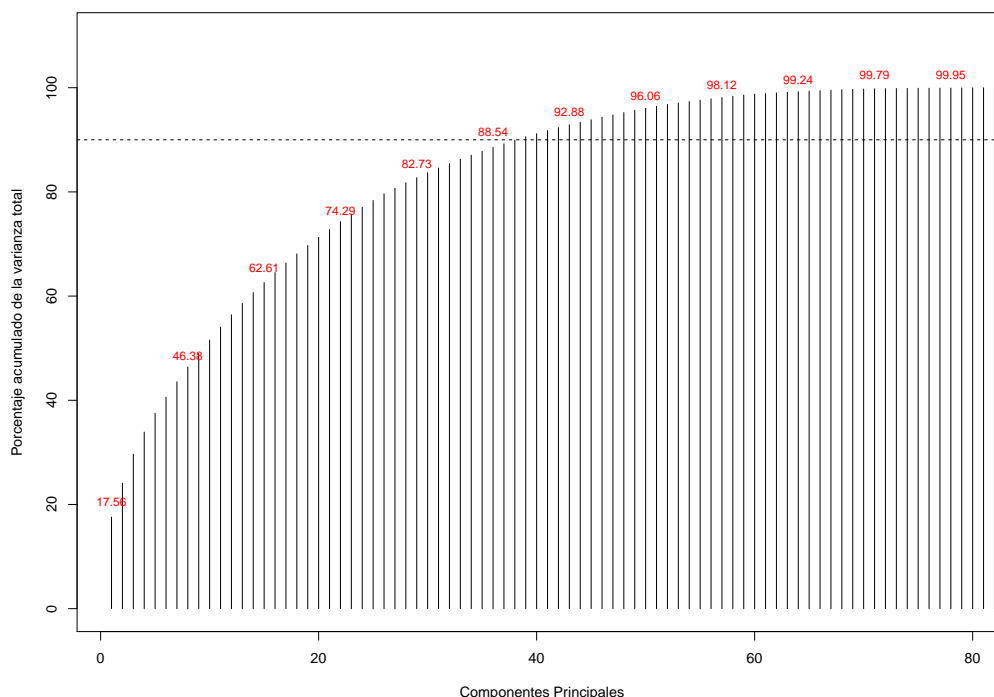


4.1.3 CNPV 2012

Se inicia con las variables del censo de 2012 descritas en el cuadro 3.1, cada una de ellas son llevadas a variables binarias. Se inicia con 84 variables, la figura 4.25 presenta un gráfico de correlaciones de estas 84 variables iniciales.

A partir de estas 84 variables se calcula el componente principal inicial, la figura 4.26 presenta el Scree plot de los eigen valores, también, se presenta en la figura 4.27 el porcentaje acumulado de la varianza.

Figura 4.27: Porcentaje acumulada de la varianza para el componente inicial para el censo de 2012

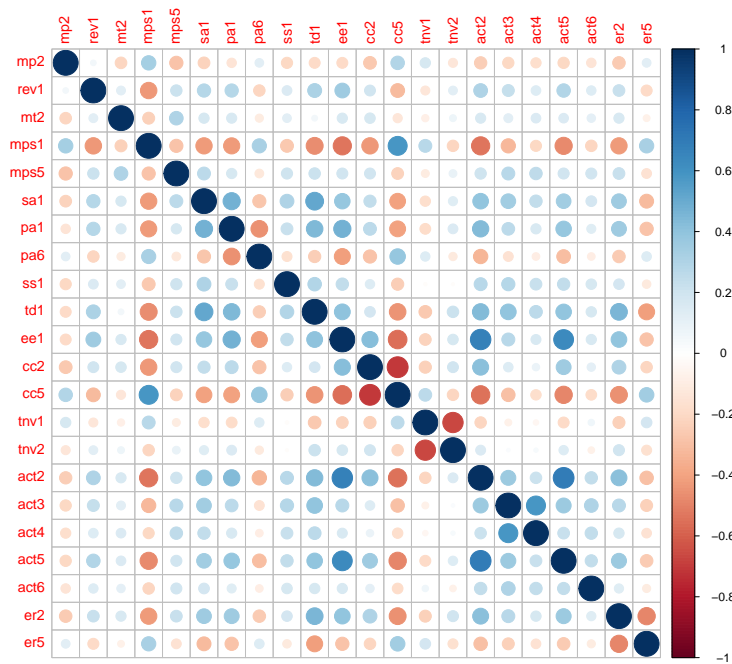


A partir de los resultados del componente inicial se siguen los siguientes pasos:

1. **En base al componente inicial se identifica el número de componentes principales que explican el 90 % de la varianza.** Se identificaron 29 componentes, es decir, los 34 componentes finales explican el 10 de la varianza.
2. **Siguiendo lo descrito en 2.5.2 se descartan 34 variables asociadas con el último componente principal,** para ello se sigue un proceso iterativo de cálculo del componente a partir de las variables relevantes y el descarte de las variables asociadas con el último componente principal.
3. Se seleccionan las variables que logren al menos una correlación absoluta superior a 30 % con el primer componente principal, del componente principal basado en las 29 variables restantes del criterio anterior (componente depurado). Finalmente, se calcula el componente principal en base a las variables que cumplen el criterio de la correlación (indicador de riqueza).

Aplicando los criterios, se terminó seleccionando un total de 22 variables. La figura 4.28 presenta las correlaciones entre las variables finales para el indicador de riqueza.

Figura 4.28: Correlaciones entre las 22 variables resultantes de los criterios, para el censo de 2012



La figura 4.29 presenta el screeplot para el componente principal final, mientras que en 4.30 se presenta el porcentaje acumulado de la varianza para el componente final.

Figura 4.29: Scree Plot del componente principal inicial para el censo de 2012

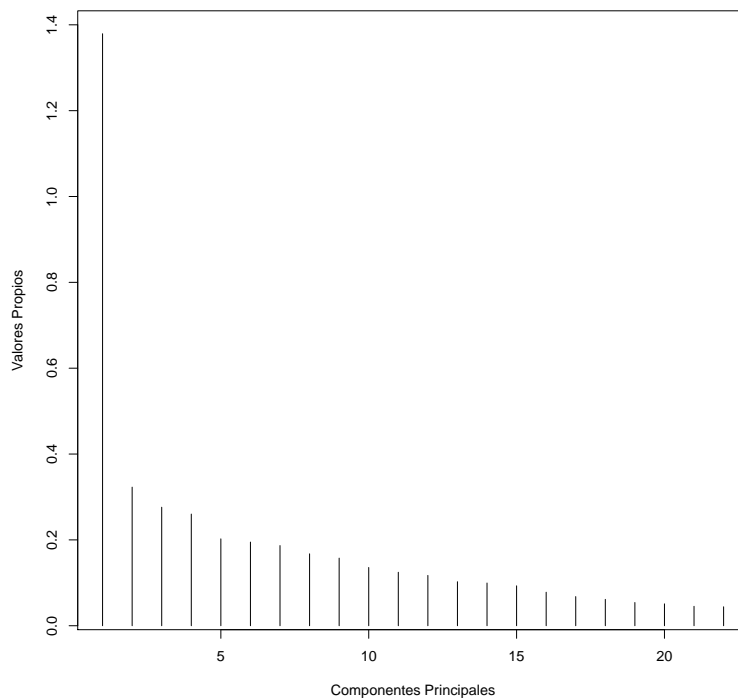
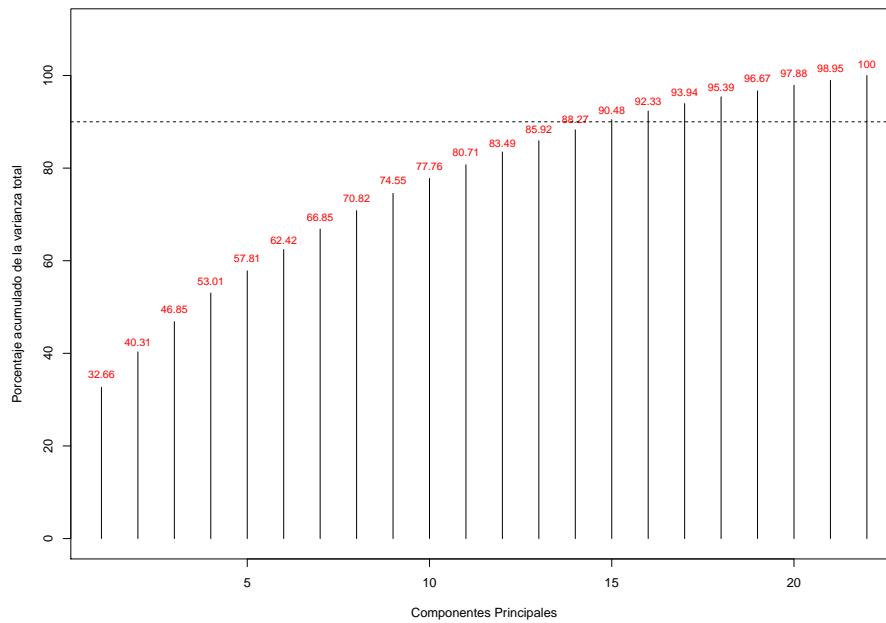
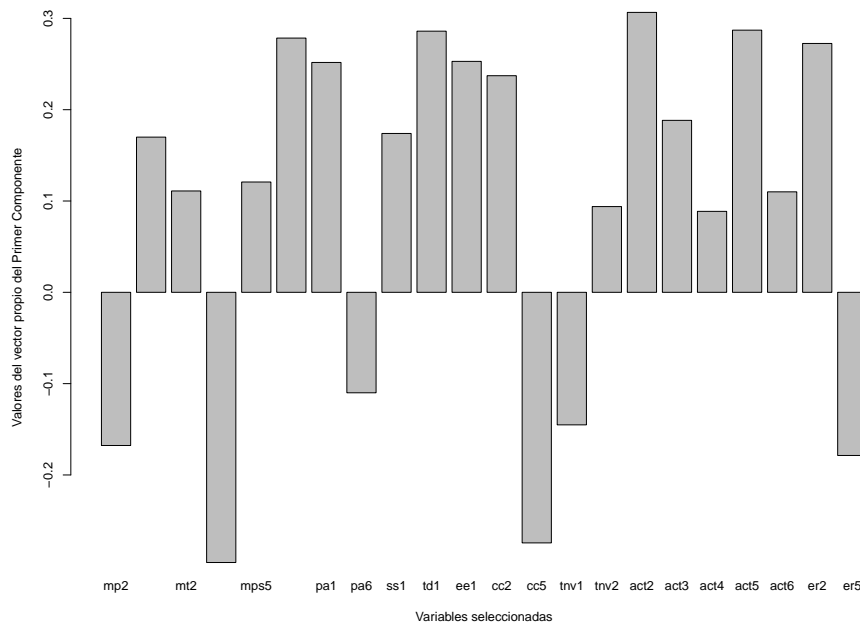


Figura 4.30: Porcentaje acumulada de la varianza para el componente inicial para el censo de 2012



La figura 4.31 presenta la dirección de las variables asociadas al primer componente según las variables finales.

Figura 4.31: Eigenvalores y dirección para el primer componente para el censo de 1992



Los cuadros 4.6 y 4.7 presentan los eigenvalores y eigenvectores respectivamente, se indica la variable asociada a cada valor dentro de los eigenvectores.

Cuadro 4.6: Eigenvalores y % acumulado para el indicador de riqueza, 2012

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}	λ_{19}	λ_{20}	λ_{21}	λ_{22}
Eigenvalores	1.38	0.32	0.28	0.26	0.20	0.19	0.19	0.17	0.16	0.14	0.12	0.12	0.10	0.10	0.09	0.08	0.07	0.06	0.05	0.05	0.05	0.04
% acumulado	32.66	40.31	46.85	53.01	57.81	62.42	66.85	70.82	74.55	77.76	80.71	83.49	85.92	88.28	90.48	92.33	93.94	95.39	96.67	97.88	98.95	100.00

Cuadro 4.7: Eigenvectores para el indicador de riqueza, 2012

	mp2	rev1	mt2	mps1	mps5	sa1	pa1	pa6	ss1	td1	ee1	cc2	cc5	tnv1	tnv2	act2	act3	act4	act5	act6	er2	er5
eigenvector 1	-0.17	0.17	0.11	-0.30	0.12	0.28	0.25	-0.11	0.17	0.29	0.25	0.24	-0.27	-0.14	0.09	0.31	0.19	0.09	0.29	0.11	0.27	-0.18
eigenvector 2	-0.03	-0.05	-0.03	-0.05	-0.09	-0.19	-0.04	-0.02	-0.39	-0.10	0.04	0.31	-0.15	-0.53	0.40	-0.01	-0.31	-0.18	-0.04	-0.30	0.10	-0.05
eigenvector 3	0.55	0.22	-0.38	0.06	-0.18	0.19	0.21	-0.02	-0.12	0.36	0.02	-0.35	0.12	-0.10	0.09	-0.03	0.00	-0.03	-0.02	-0.14	0.11	-0.22
eigenvector 4	0.39	0.16	-0.17	-0.04	-0.18	-0.22	0.04	-0.09	-0.14	-0.24	0.27	0.29	-0.18	0.31	-0.23	0.30	-0.13	-0.11	0.33	0.05	-0.18	0.15
eigenvector 5	0.21	0.34	0.37	-0.07	0.12	0.11	0.13	-0.05	0.18	-0.10	0.07	-0.14	0.15	-0.32	0.24	0.01	0.01	0.04	0.05	-0.03	-0.50	0.38
eigenvector 6	-0.21	-0.28	-0.58	0.08	-0.19	-0.05	-0.11	0.02	0.46	0.12	0.05	-0.02	-0.03	-0.23	0.13	0.15	0.06	0.01	0.17	-0.06	-0.27	0.22
eigenvector 7	0.02	-0.03	0.01	-0.02	-0.06	0.33	0.31	-0.12	0.39	-0.01	-0.02	0.27	-0.16	0.19	-0.14	-0.21	-0.38	-0.21	-0.33	-0.34	-0.04	0.04
eigenvector 8	-0.24	-0.40	0.20	0.13	-0.12	0.14	0.39	-0.10	-0.24	-0.04	0.19	-0.36	0.23	0.07	-0.05	0.24	-0.19	-0.12	0.28	-0.18	0.03	0.10
eigenvector 9	-0.19	-0.02	-0.28	-0.04	0.03	0.42	0.18	-0.05	-0.55	0.08	-0.10	0.18	-0.13	0.04	-0.03	-0.12	0.17	0.09	-0.15	0.22	-0.39	0.17
eigenvector 10	0.43	-0.48	0.20	0.44	-0.06	0.09	0.17	-0.05	0.09	-0.13	-0.06	0.30	-0.08	-0.15	0.09	-0.02	0.27	0.19	0.01	0.16	0.03	-0.09
eigenvector 11	0.01	0.01	0.08	0.04	0.07	0.07	-0.21	0.07	-0.11	0.06	0.01	0.09	-0.05	0.12	-0.12	0.05	0.42	0.22	0.10	-0.79	-0.09	0.01
eigenvector 12	0.13	-0.10	0.10	0.07	0.04	0.58	-0.59	0.22	-0.01	0.03	0.02	0.02	0.02	0.16	-0.30	-0.17	0.19	0.09	-0.10	-0.15	-0.01	-0.15
eigenvector 13	-0.00	0.18	-0.32	0.16	0.27	0.30	-0.01	-0.07	0.02	-0.55	0.03	-0.05	0.09	-0.06	0.06	-0.02	0.06	0.05	0.06	-0.04	0.49	0.32
eigenvector 14	-0.12	0.08	-0.17	0.00	0.18	0.01	0.16	-0.02	0.08	-0.46	0.06	-0.04	0.12	-0.02	0.04	0.02	0.02	0.04	0.06	-0.05	-0.36	-0.72
eigenvector 15	-0.05	0.04	0.12	-0.34	-0.79	0.20	-0.08	0.03	0.04	-0.34	-0.05	-0.07	0.03	-0.08	-0.02	-0.01	0.22	0.07	-0.10	0.03	0.08	-0.04
eigenvector 16	0.32	-0.49	-0.09	-0.70	0.32	-0.01	-0.02	-0.02	-0.01	-0.12	0.02	-0.13	0.01	-0.03	-0.03	0.04	0.04	-0.01	-0.12	-0.03	0.00	0.06
eigenvector 17	-0.02	0.03	0.00	0.16	0.02	-0.01	-0.12	-0.11	-0.02	-0.01	0.34	-0.08	-0.01	0.01	0.04	0.59	0.05	0.04	-0.69	0.01	-0.02	-0.01
eigenvector 18	-0.02	-0.07	0.04	0.06	-0.03	0.02	-0.27	-0.52	-0.02	0.01	0.57	-0.19	-0.20	0.01	0.03	-0.48	0.01	0.09	0.04	0.04	-0.04	-0.04
eigenvector 19	0.01	-0.04	-0.01	-0.06	-0.05	0.00	0.04	0.17	0.01	-0.01	0.08	0.02	-0.00	0.57	0.77	-0.04	-0.05	0.18	0.01	-0.01	0.01	0.03
eigenvector 20	-0.01	-0.02	-0.04	-0.00	-0.01	-0.01	0.14	0.67	-0.00	0.02	0.59	0.14	0.19	-0.12	-0.17	-0.25	-0.01	0.09	-0.08	0.04	0.04	0.04
eigenvector 21	-0.00	-0.01	0.05	0.09	0.05	-0.01	0.06	0.23	0.01	-0.10	0.05	-0.23	-0.42	0.05	0.11	-0.06	0.40	-0.71	-0.02	0.01	-0.03	0.01
eigenvector 22	-0.01	-0.01	-0.02	-0.06	-0.01	0.01	-0.08	-0.27	0.01	0.12	0.10	0.38	0.67	0.09	0.10	-0.05	0.28	-0.45	-0.01	0.02	0.03	-0.02

A partir de los valores dados por el primer componente resultante de estas 22 variables, se calcula el indicador de riqueza para el censo 2012, la figura 4.32 presenta los histogramas y la densidad del indicador de riqueza, en las figuras 4.33, 4.34 y 4.35 presentan los diagramas de caja para el indicador de riqueza global, por área y departamento para el indicador obtenido para el 2012..

Figura 4.32: Densidad e histograma a partir del indicador para 2012

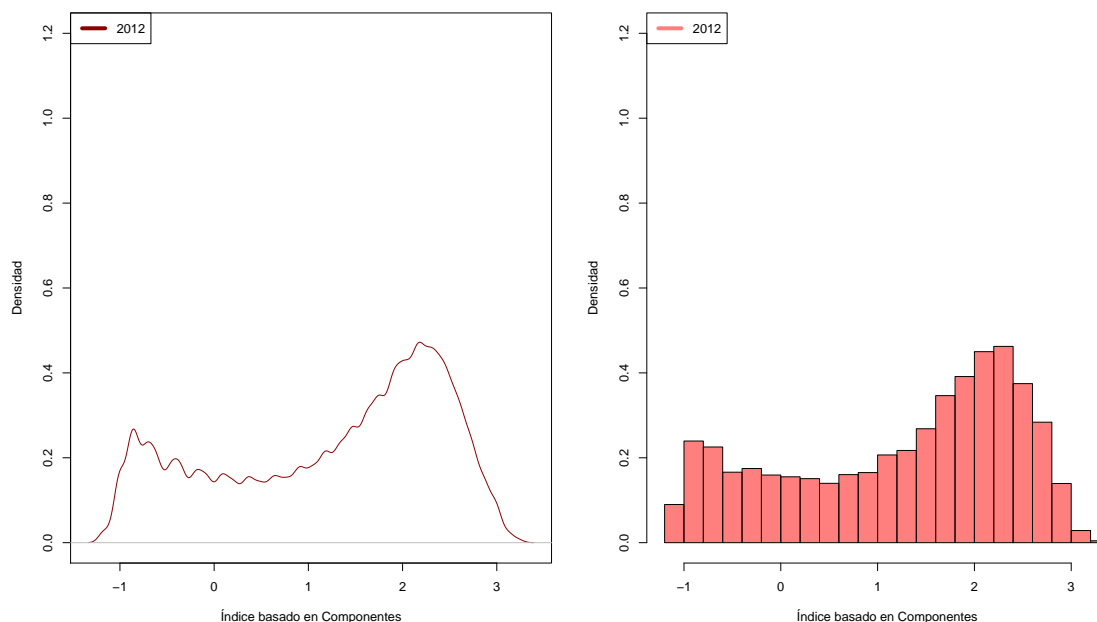


Figura 4.33: Diagrama de caja del indicador de riqueza, Censo 2012

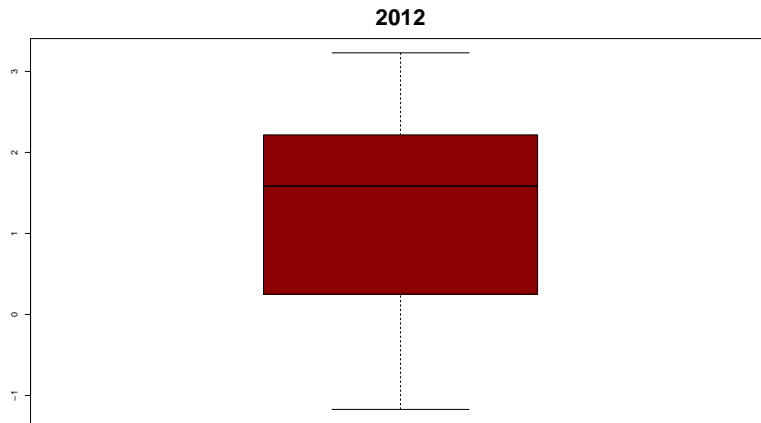


Figura 4.34: Diagrama de caja por área del indicador de riqueza, Censo 2012

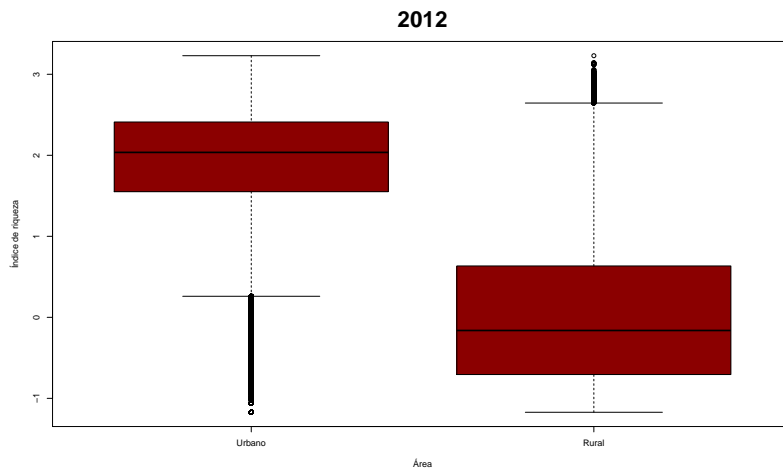
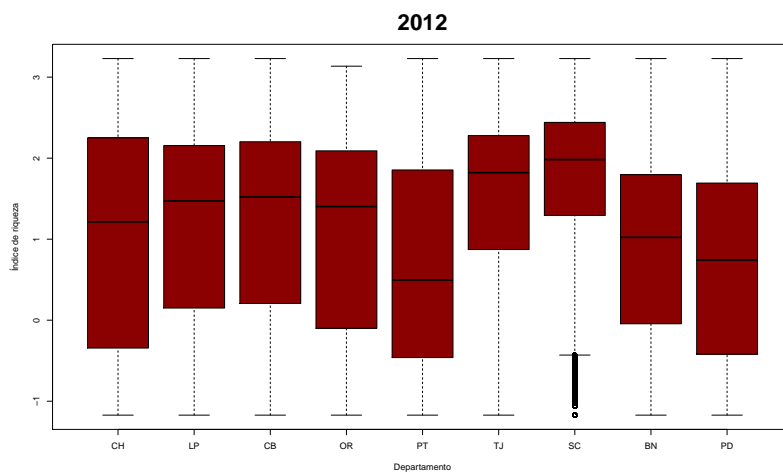
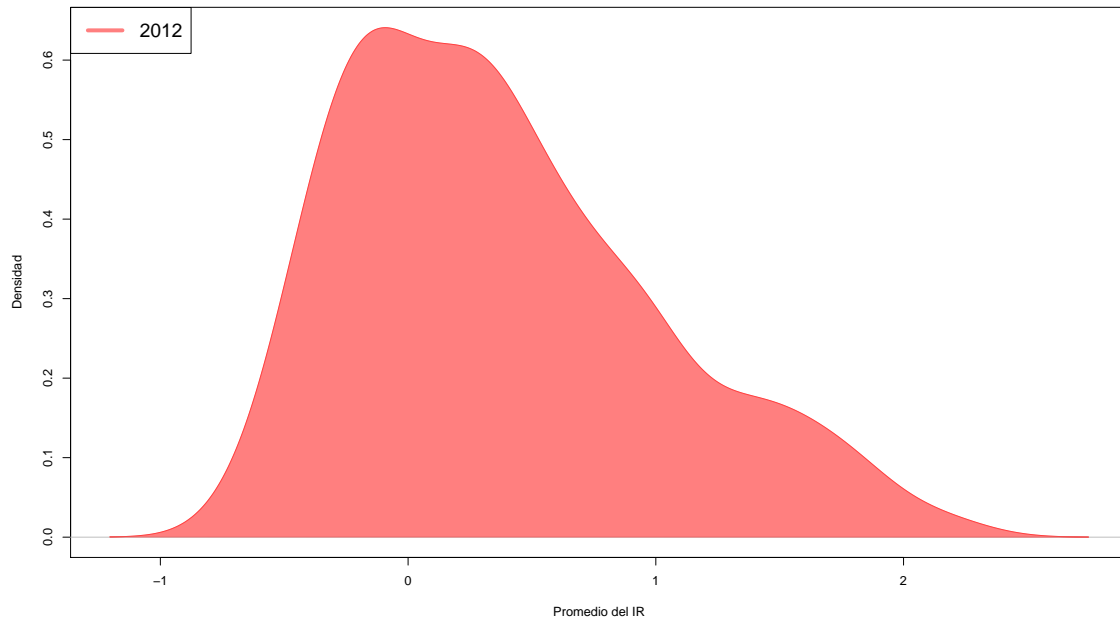


Figura 4.35: Diagrama de caja por departamento del indicador de riqueza, Censo 2012



A partir del indicador de riqueza se obtiene el promedio del indicador a nivel municipal, la figura 4.36 presenta la densidad del indicador a nivel municipal para el censo de 2012.

Figura 4.36: Indicador de riqueza a nivel municipal en base al censo de 2012

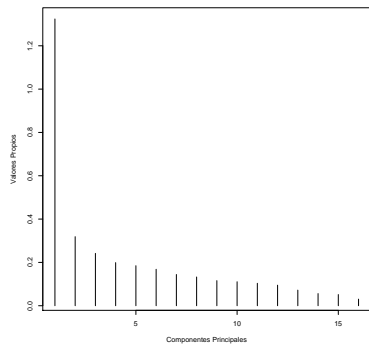


4.1.4 Para el CNPV 2012, a nivel de los 9 departamentos

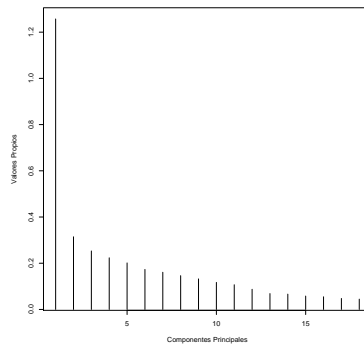
En los resultados anteriores se definió el indicador de riqueza en base al año del censo de interés y se calculó el indicador para los censos posteriores, esto permite ver los cambios en la riqueza de forma dinámica en tiempo condicionado al año de procedencia del indicador. Sin embargo, para el caso del censo 2012 no es posible ver los avances hacia adelante debido a que este es el último censo, se podría visualizar los avances empleando encuestas nacionales que compartan un set similar de preguntas con el censo 2012, esto escapa del alcance de este trabajo. Una alternativa para explotar las ventajas de la metodología sugerida es comparar dominios dentro de un mismo periodo, definiendo un dominio base para la medición del indicador de riqueza.

Se calculó el indicador de riqueza para cada departamento, siguiendo el mismo procedimiento descrito en las anteriores secciones, dada la definición del indicador para cada departamento se procedió a calcular el indicador para el resto de los departamentos. En la figura 4.37 se presentan las correlaciones de las variables finales para cada departamento, en 4.38 se muestra el scree plot para cada departamento, en 4.39 se muestra el porcentaje acumulado de la varianza y finalmente en 4.40 se muestran las direcciones de las variables para el primer componente de cada departamento.

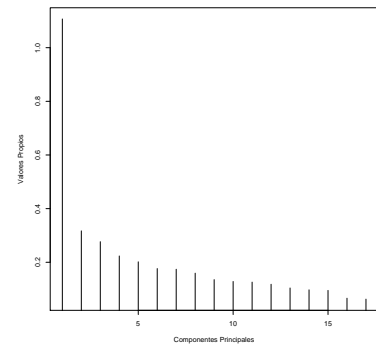
Figura 4.38: Scree Plot del componente principal para el indicador de riqueza, por departamento. Censo 2012



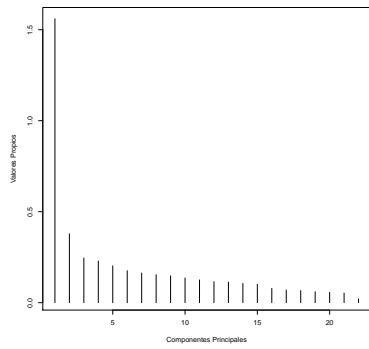
(a) Chuquisaca



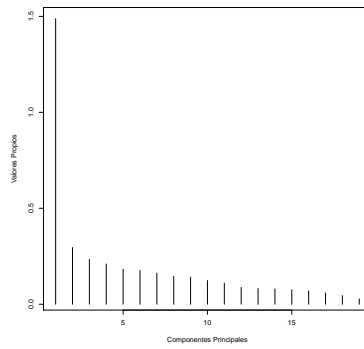
(b) La Paz



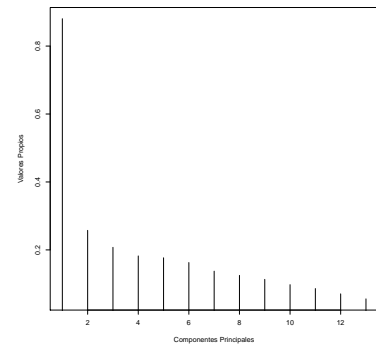
(c) Cochabamba



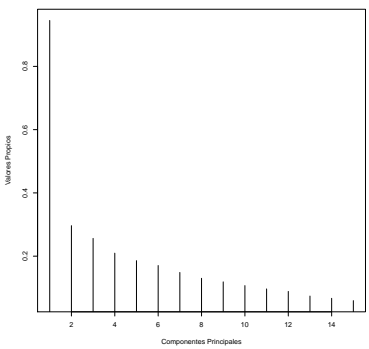
(d) Oruro



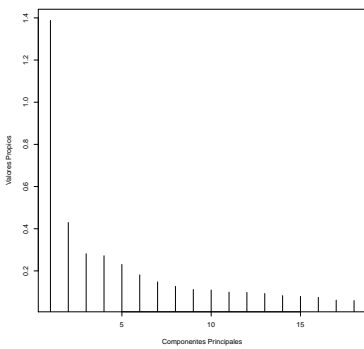
(e) Potosí



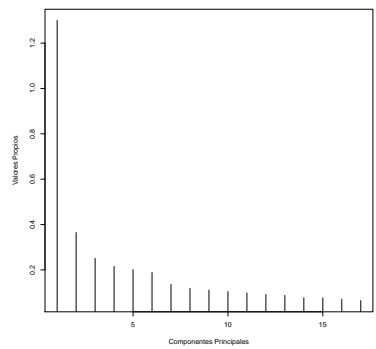
(f) Tarija



(g) Santa Cruz

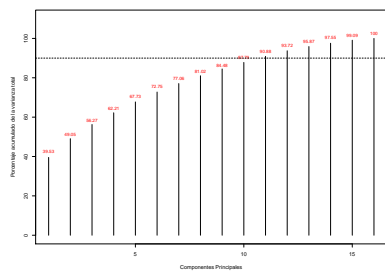


(h) Beni

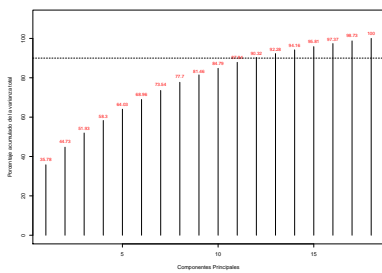


(i) Pando

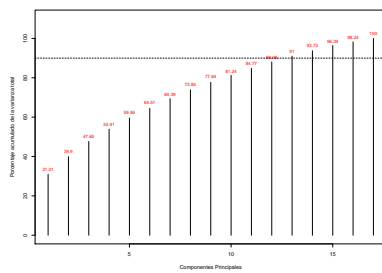
Figura 4.39: Porcentaje acumulado de la varianza para el componente principal para el indicador de riqueza, por departamento. Censo 2012



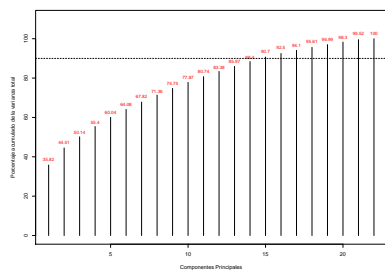
(a) Chuquisaca



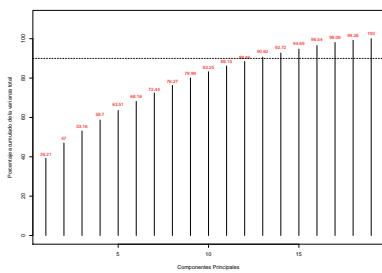
(b) La Paz



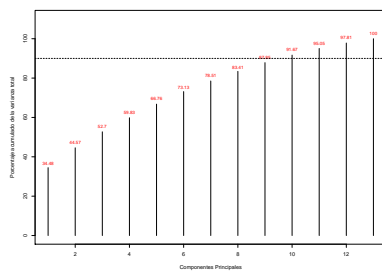
(c) Cochabamba



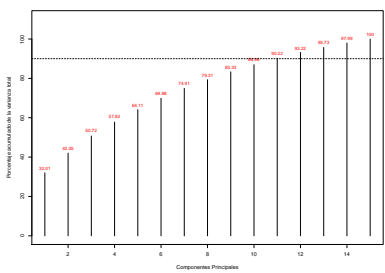
(d) Oruro



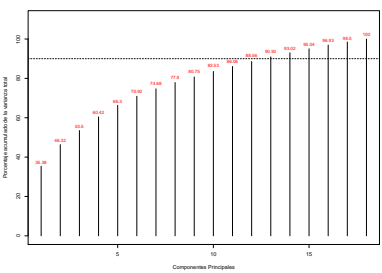
(e) Potosí



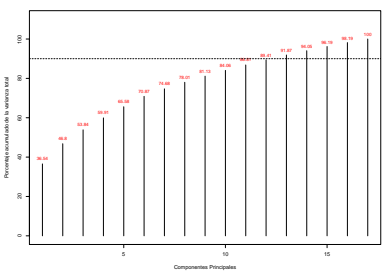
(f) Tarija



(g) Santa Cruz

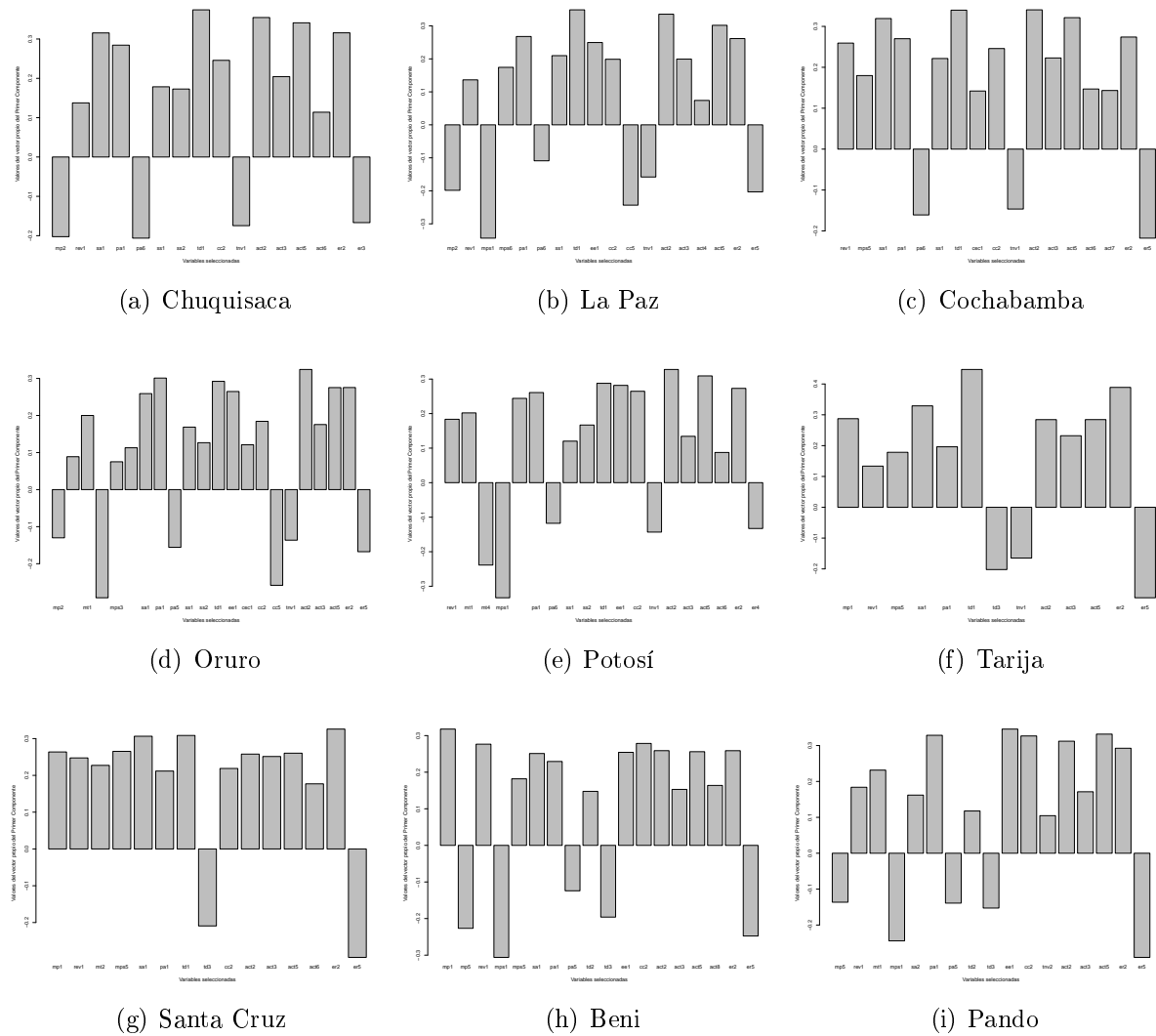


(h) Beni



(i) Pando

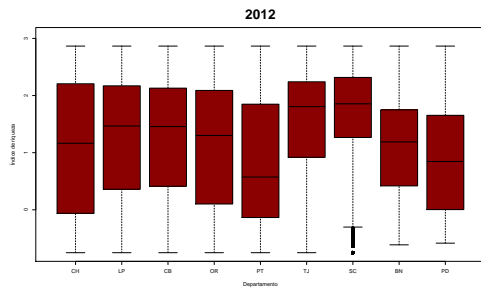
Figura 4.40: Eigenvalores y dirección para el primer componente del indicador de riqueza, por departamento. Censo 2012



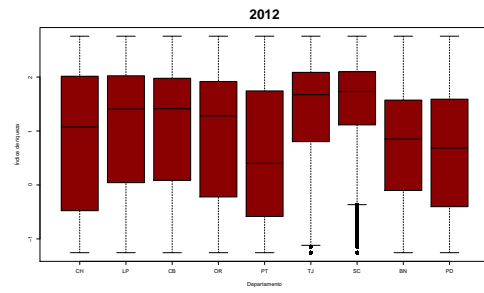
Los resultados de los eigenvalores y eigenvectores finales se presentan en el anexo A.2.

A partir de estos resultados se aplicaron los valores del indicador para cada departamento según el departamento base. En la figura 4.41 se presentan los gráficos de caja para los nueve departamentos

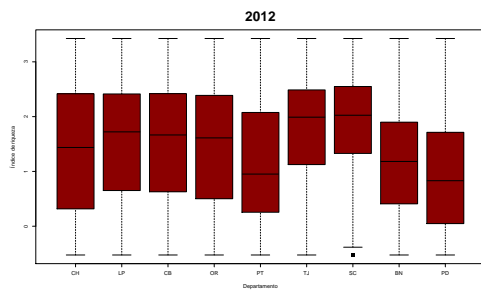
Figura 4.41: Diagramas de caja según el indicador de riqueza, base a nivel departamental



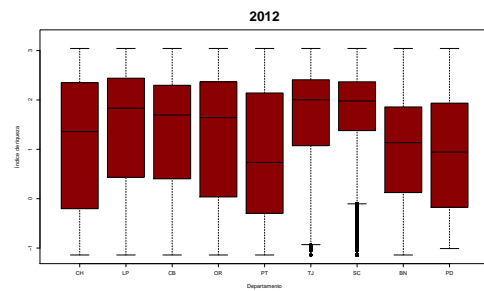
(a) Chuquisaca



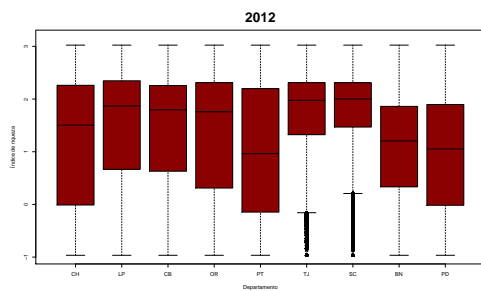
(b) La Paz



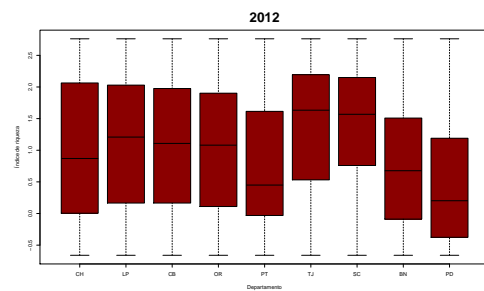
(c) Cochabamba



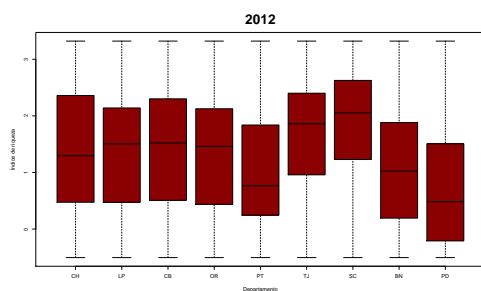
(d) Oruro



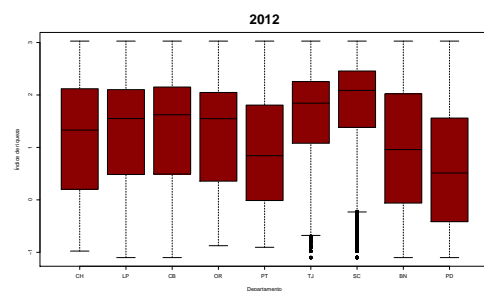
(e) Potosí



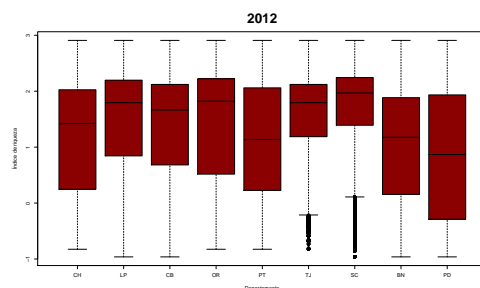
(f) Tarija



(g) Santa Cruz



(h) Beni



(i) Pando

4.2 Conclusiones, recomendaciones y limitaciones

Las conclusiones del trabajo son:

- Se logró cumplir los objetivos del trabajo:
 - Se realizó la revisión de la literatura
 - Se logró recopilar la información para el trabajo
 - Se adecuó la información de las bases de datos recolectadas
 - Se estimó y ajustó el modelo a partir de la metodología propuesta
 - Se obtuvieron los resultados para los dominios y los tres censos.
- Se presentó la estructura de los indicadores de riqueza según la base del censo, presentado en el cuadro 4.1, en este se evidencia los cambios en el tiempo respecto a las variables relevantes y las ponderaciones correspondientes
- A partir de los resultados se evidencia que en todos los casos el departamento de Tarija y Santa Cruz muestran los mayores avances, tanto a nivel del indicador y su avance en el tiempo.
- Para el censo de 2012 se construyó el indicador en base a cada uno de los departamentos, a partir de esto se puede analizar el bienestar definido a partir de un departamento y compararlo con el resto de departamentos.

Las recomendaciones:

- Es importante tener en cuenta que el indicador de riqueza está a nivel de un censo de población y permite monitorear el bienestar a partir del censo que se use como base.
- El indicador de riqueza se construye a partir de variables binarias provenientes de variables relacionadas del hogar y la vivienda, también se incluyen a variables relacionadas con dominios de estudio como el área o el departamento.

Es importante mencionar limitaciones del indicador propuesto:

- El indicador no permite una interpretación completa de cambios en el tiempo, pues indica un cambio relativo respecto a los demás hogares pero no indica si el hogar en sí mismo ha mejorado respecto a una situación anterior.
- Al no considerar la inclusión de variables referidas a educación y salud que viene a nivel de persona, limita su comparabilidad con el NBI.

Capítulo 5

Propuesta de Mejoramiento

A partir de este trabajo se puede mejorar el planteamiento para el indicador en base a los siguientes puntos:

1. Se puede explorar la inclusión de otras variables relacionadas a los miembros del hogar, como educación, empleo, salud etc. Si bien el método propuesto considera un paso de filtración de las variables iniciales, es importante discutir la pertinencia de que variables incluir en la fase inicial, pensando que este indicador puede emplearse como un indicador de monitoreo a partir de estudios posteriores con un soporte común.
2. Se puede evaluar la utilidad de este indicador para otros mecanismos estadísticos, como variables para estratificación.
3. Existe alternativas para componentes principales con variables binarias, sin embargo, estas alternativas no permiten tener una interpretación amigable y sencilla para ser usada en poblaciones sin una formación técnica. En próximos trabajos se puede comparar el método propuesto con las otras alternativas.
4. Para aplicar el método a partir de una muestra probabilística es importante entender el comportamiento y las propiedades de los distribuciones muestrales, a partir de esto, calcular el error de muestreo.
5. A partir del indicador dado se puede pensar en transformaciones que permitan una interpretación mas amigable y posterior toma de decisiones; se pueden usar transformaciones vinculadas a los cuantiles de la distribución o opciones para relativizarlo.
6. Dado que los censos contienen múltiples divisiones geográficas, es importante hacia adelante pensar en alternativas para la agregación de la información proveniente del indicador propuesto.

Apéndice A

Anexos

A.1 Siglas

- *CNPV* : Censo Nacional de Población y Vivienda
- *NBI* : Índice de Necesidades Básicas Insatisfechas
- *MA* : Ministerio de Autonomías
- *PCA* : Análisis de componentes principales
- *UNESCO* : Organización de las Naciones Unidas para la Educación
- *MANOVA* : Análisis multivariante de la varianza
- *FA* : Análisis Factorial
- *PDES* : Plan de Desarrollo Económico y Social
- *CPs* : Componentes Principales
- *SJ* : Satisfacción con el trabajo
- *SJT* : Satisfacción con la capacitación en el trabajo
- *SWC* : Satisfacción con las condiciones de trabajo
- *SMC* : Satisfacción con el seguro medico
- *SDC* : Satisfacción con el seguro dental
- *PVT* : Proporción de la varianza total explicada

A.2 Eigenvalores y eigenvectores finales para el censo 2012 por departamentos

Cuadro A.1: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Chuquisaca

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}
Eigenvalores	1.32	0.32	0.24	0.20	0.18	0.17	0.14	0.13	0.12	0.11	0.10	0.09	0.07	0.06	0.05	0.03
% acumulado	39.53	49.05	56.27	62.21	67.73	72.75	77.06	81.02	84.48	87.79	90.88	93.72	95.87	97.55	99.09	100.00

Cuadro A.2: Eigenvalores y % acumulado para el indicador de riqueza, 2012. La Paz

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}
Eigenvalores	1.26	0.31	0.25	0.22	0.20	0.17	0.16	0.15	0.13	0.12	0.11	0.09	0.07	0.07	0.06	0.05	0.05	0.04
% acumulado	35.78	44.73	51.93	58.30	64.03	68.95	73.54	77.70	81.46	84.79	87.84	90.32	92.28	94.16	95.81	97.38	98.73	100.00

Cuadro A.3: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Cochabamba

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}
Eigenvalores	1.11	0.32	0.28	0.22	0.20	0.18	0.17	0.16	0.14	0.13	0.13	0.12	0.10	0.10	0.10	0.07	0.06
% acumulado	31.01	39.90	47.65	53.91	59.56	64.51	69.39	73.85	77.64	81.24	84.78	88.08	91.00	93.73	96.39	98.25	100.00

Cuadro A.4: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Oruro

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}	λ_{19}	λ_{20}	λ_{21}	λ_{22}
Eigenvalores	1.56	0.38	0.25	0.23	0.20	0.18	0.16	0.15	0.15	0.14	0.13	0.12	0.11	0.11	0.10	0.08	0.07	0.07	0.06	0.06	0.05	0.02
% acumulado	35.82	44.51	50.14	55.40	60.04	64.08	67.82	71.36	74.75	77.87	80.74	83.38	85.97	88.40	90.70	92.50	94.10	95.61	96.99	98.30	99.52	100.00

Cuadro A.5: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Potosi

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}	λ_{19}	
Eigenvalores	1.49	0.30	0.23	0.21	0.18	0.18	0.16	0.15	0.14	0.12	0.11	0.09	0.08	0.08	0.07	0.07	0.06	0.06	0.04	0.03
% acumulado	39.21	47.00	53.16	58.70	63.51	68.16	72.44	76.27	79.99	83.25	86.15	88.46	90.62	92.72	94.69	96.54	98.09	99.26	100.00	

Cuadro A.6: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Tarija

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}
Eigenvalores	0.88	0.26	0.21	0.18	0.18	0.16	0.14	0.12	0.11	0.10	0.09	0.07	0.06
% acumulado	34.48	44.57	52.70	59.83	66.76	73.13	78.51	83.41	87.84	91.67	95.05	97.81	100.00

Cuadro A.7: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Santa Cruz

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}
Eigenvalores	0.95	0.30	0.26	0.21	0.19	0.17	0.15	0.13	0.12	0.11	0.10	0.09	0.07	0.07	0.06
% acumulado	32.01	42.05	50.72	57.81	64.11	69.88	74.91	79.31	83.33	86.96	90.22	93.22	95.72	97.99	100.00

Cuadro A.8: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Beni

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}	λ_{18}
Eigenvalores	1.39	0.43	0.28	0.27	0.23	0.18	0.15	0.13	0.11	0.11	0.10	0.10	0.09	0.08	0.08	0.07	0.06	0.06
% acumulado	35.38	46.33	53.50	60.42	66.30	70.92	74.68	77.90	80.75	83.53	86.06	88.56	90.92	93.02	95.04	96.93	98.50	100.00

Cuadro A.9: Eigenvalores y % acumulado para el indicador de riqueza, 2012. Pando

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}
Eigenvalores	1.30	0.36	0.25	0.22	0.20	0.19	0.14	0.12	0.11	0.10	0.10	0.09	0.09	0.08	0.08	0.07	0.06
% acumulado	36.54	46.80	53.84	59.91	65.58	70.87	74.68	78.01	81.13	84.06	86.81	89.41	91.87	94.05	96.19	98.19	100.00

Cuadro A.10: Eigenvectores para el indicador de riqueza, 2012. Chuquisaca

	mp2	rev1	sal	pa1	pa6	ss1	ss2	td1	cc2	tnv1	act2	act3	act5	act6	er2	er3
eigenvector 1	-0.20	0.14	0.32	0.28	-0.21	0.18	0.17	0.37	0.25	-0.17	0.35	0.20	0.34	0.11	0.32	-0.17
eigenvector 2	-0.03	-0.08	-0.12	-0.07	0.02	-0.61	0.58	0.02	0.28	-0.34	-0.04	-0.17	-0.04	-0.18	0.05	-0.04
eigenvector 3	0.72	0.32	0.22	0.40	-0.28	-0.08	0.10	-0.09	-0.02	0.07	-0.01	-0.14	0.01	-0.12	-0.10	0.11
eigenvector 4	0.12	-0.08	-0.21	-0.10	-0.06	0.26	-0.32	-0.18	0.71	-0.22	0.17	-0.23	0.17	-0.12	-0.22	0.03
eigenvector 5	-0.26	-0.12	0.30	0.28	-0.18	0.20	-0.15	0.07	0.02	-0.33	-0.40	-0.20	-0.45	-0.36	0.07	-0.03
eigenvector 6	0.38	0.35	-0.15	-0.32	0.33	0.18	-0.09	0.18	-0.08	-0.40	-0.10	0.16	-0.09	-0.16	0.36	-0.24
eigenvector 7	-0.29	0.54	-0.08	0.01	0.01	0.01	0.02	-0.04	-0.07	-0.33	-0.02	0.18	0.03	0.05	-0.26	0.64
eigenvector 8	0.27	-0.62	0.10	0.01	0.09	0.03	-0.04	0.12	-0.22	-0.29	0.14	0.17	0.20	-0.15	0.11	0.50
eigenvector 9	0.13	-0.02	0.41	-0.22	0.25	-0.00	0.06	0.17	0.47	0.25	-0.21	0.08	-0.31	0.33	0.15	0.33
eigenvector 10	-0.15	0.18	0.20	-0.20	0.23	0.04	0.04	0.13	-0.02	0.40	0.18	-0.30	0.20	-0.66	0.10	0.18
eigenvector 11	-0.01	-0.01	-0.64	0.25	-0.20	0.10	0.04	0.19	0.10	0.26	-0.07	-0.09	-0.14	0.00	0.49	0.29
eigenvector 12	0.01	-0.04	-0.07	0.06	-0.10	-0.10	-0.01	-0.10	0.23	0.21	-0.10	0.79	-0.08	-0.44	-0.11	-0.08
eigenvector 13	-0.10	0.09	0.19	-0.06	-0.12	-0.33	-0.34	-0.58	0.02	-0.05	-0.03	0.02	0.17	0.03	0.57	0.08
eigenvector 14	0.00	0.02	0.02	-0.06	-0.06	-0.07	-0.06	-0.08	-0.06	-0.05	0.74	0.03	-0.64	-0.02	0.03	0.03
eigenvector 15	-0.05	0.00	-0.06	0.64	0.73	-0.07	-0.07	-0.15	0.09	-0.01	0.08	0.00	-0.01	0.02	0.01	-0.05
eigenvector 16	0.01	0.06	-0.03	0.05	-0.05	-0.56	-0.60	0.55	0.02	0.01	-0.00	-0.04	0.03	0.01	-0.12	-0.02

Cuadro A.11: Eigenvectores para el indicador de riqueza, 2012. La Paz

	mp2	rev1	mps1	mps6	pal	pa6	ssl	td1	ee1	cc2	cc5	tnv1	act2	act3	act4	act5	er2	er5
eigenvector 1	-0.20	0.14	-0.34	0.17	0.27	-0.11	0.21	0.35	0.25	0.20	-0.24	-0.16	0.34	0.20	0.07	0.30	0.26	-0.20
eigenvector 2	-0.20	-0.05	0.17	-0.54	0.07	0.03	0.41	0.24	-0.07	-0.45	0.19	0.07	-0.03	0.34	0.18	-0.02	0.12	-0.02
eigenvector 3	0.70	0.33	0.07	-0.04	0.18	-0.14	0.04	-0.01	0.27	-0.27	0.10	0.16	0.20	-0.01	-0.02	0.29	-0.08	0.18
eigenvector 4	-0.00	-0.05	-0.22	0.51	0.08	0.03	-0.14	0.17	-0.10	-0.62	0.20	0.09	-0.10	-0.12	-0.09	-0.16	0.32	-0.15
eigenvector 5	0.34	-0.01	0.20	-0.33	0.15	-0.08	-0.33	0.20	-0.08	0.11	-0.09	-0.45	-0.12	-0.11	-0.04	-0.17	0.46	-0.23
eigenvector 6	-0.28	-0.25	0.12	-0.14	-0.07	0.08	-0.61	-0.06	0.20	-0.23	0.15	-0.12	0.33	0.04	-0.01	0.45	-0.06	-0.00
eigenvector 7	0.02	-0.27	0.15	-0.10	0.09	-0.07	-0.06	-0.13	0.04	0.17	-0.18	0.76	0.06	-0.06	-0.03	0.06	0.43	-0.15
eigenvector 8	0.13	0.38	-0.15	0.03	-0.50	0.19	-0.29	0.00	-0.16	0.05	-0.09	0.14	-0.02	0.54	0.24	-0.02	0.19	-0.01
eigenvector 9	-0.18	0.22	-0.06	-0.06	0.56	-0.30	-0.42	0.06	-0.03	0.01	-0.04	0.20	-0.18	0.25	0.11	-0.25	-0.32	0.03
eigenvector 10	-0.41	0.58	0.01	-0.18	-0.08	-0.06	-0.04	-0.08	0.12	-0.06	-0.03	0.03	-0.05	-0.39	-0.15	0.00	0.33	0.37
eigenvector 11	-0.04	0.32	0.00	-0.18	-0.18	0.12	-0.02	0.18	0.03	-0.07	-0.05	0.18	0.06	-0.29	-0.13	-0.01	-0.34	-0.72
eigenvector 12	-0.07	0.19	0.03	0.06	0.18	-0.07	0.12	-0.71	0.04	0.03	0.33	-0.17	-0.14	0.14	0.12	0.14	0.17	-0.40
eigenvector 13	-0.02	-0.07	0.06	0.06	-0.14	-0.21	0.01	0.06	0.13	-0.17	-0.38	-0.04	-0.70	0.02	0.03	0.47	-0.05	-0.04
eigenvector 14	0.02	0.08	-0.05	0.01	0.19	0.32	-0.04	0.32	-0.25	0.36	0.53	0.13	-0.29	-0.08	-0.04	0.40	0.02	0.07
eigenvector 15	0.03	-0.12	-0.16	-0.07	-0.13	0.11	-0.04	0.11	0.80	0.13	0.29	0.04	-0.26	0.02	0.11	-0.30	0.03	-0.01
eigenvector 16	0.13	-0.16	-0.80	-0.43	0.05	0.04	-0.04	-0.21	-0.14	-0.10	-0.08	0.01	-0.05	-0.19	0.09	0.08	-0.01	0.04
eigenvector 17	0.01	0.06	0.10	0.02	0.37	0.79	0.00	-0.15	0.13	-0.13	-0.39	-0.03	-0.08	0.06	-0.05	-0.03	-0.01	0.05
eigenvector 18	0.00	-0.02	-0.14	-0.11	-0.03	-0.05	0.02	-0.05	0.04	0.01	0.04	-0.00	-0.06	0.39	-0.89	-0.02	0.01	-0.01

Cuadro A.12: Eigenvectores para el indicador de riqueza, 2012. Cochabamba

	rev1	mps5	sal	pal	pa6	ssl	td1	cec1	cc2	tnv1	act2	act3	act5	act6	act7	er2	er5
eigenvector 1	0.26	0.18	0.32	0.27	-0.16	0.22	0.34	0.14	0.25	-0.15	0.34	0.22	0.32	0.15	0.14	0.27	-0.22
eigenvector 2	-0.04	0.06	0.09	0.01	0.04	0.35	-0.10	0.34	-0.24	0.46	0.03	0.20	0.05	0.31	0.35	-0.37	0.27
eigenvector 3	-0.01	0.14	0.37	0.34	0.07	0.15	0.32	0.04	-0.43	-0.03	-0.33	0.08	-0.34	-0.10	-0.36	-0.10	-0.14
eigenvector 4	-0.08	0.16	-0.17	-0.52	0.18	0.06	0.10	0.13	-0.17	0.19	-0.13	0.20	-0.13	0.24	0.01	0.45	-0.46
eigenvector 5	0.02	0.08	-0.12	-0.22	-0.01	0.59	-0.12	0.13	0.25	-0.02	0.09	-0.04	0.10	-0.03	-0.66	-0.07	0.15
eigenvector 6	0.28	0.20	-0.01	-0.11	0.05	-0.57	0.01	0.07	-0.11	-0.05	0.11	0.27	0.19	0.39	-0.41	-0.26	0.11
eigenvector 7	0.22	-0.22	-0.01	0.25	-0.09	-0.24	-0.14	0.56	0.10	0.48	0.00	-0.26	-0.06	-0.14	-0.20	0.23	-0.13
eigenvector 8	0.42	0.15	-0.02	-0.33	0.12	0.01	0.02	0.44	-0.14	-0.38	-0.03	-0.02	-0.05	-0.48	0.23	-0.17	0.03
eigenvector 9	0.35	0.36	0.29	-0.13	-0.06	-0.02	-0.14	-0.20	0.49	0.17	-0.29	-0.06	-0.43	0.14	0.11	0.00	0.11
eigenvector 10	-0.63	0.04	0.32	-0.12	0.02	-0.20	0.06	0.47	0.32	-0.27	-0.05	-0.00	-0.12	0.11	-0.00	-0.02	0.12
eigenvector 11	-0.17	0.16	0.42	-0.28	0.05	-0.14	0.04	-0.21	-0.02	0.44	0.19	0.13	0.27	-0.54	-0.08	-0.02	0.02
eigenvector 12	0.20	-0.55	0.51	-0.34	0.08	0.07	0.00	-0.09	-0.12	-0.08	0.08	-0.35	0.04	0.28	-0.03	-0.08	-0.12
eigenvector 13	0.08	-0.55	-0.06	-0.02	0.03	0.01	0.11	-0.00	0.33	0.05	-0.12	0.66	-0.17	-0.15	-0.03	-0.21	-0.14
eigenvector 14	-0.11	0.19	-0.09	0.06	0.01	0.01	-0.07	-0.00	0.18	0.05	0.06	-0.22	0.06	0.01	0.03	-0.58	-0.71
eigenvector 15	-0.02	0.03	0.26	0.11	-0.13	0.04	-0.81	0.01	-0.18	-0.22	-0.03	0.30	0.05	-0.03	-0.03	0.15	-0.20
eigenvector 16	0.04	0.01	0.05	0.24	0.91	0.02	-0.11	-0.02	0.17	-0.02	-0.12	-0.02	0.21	0.01	0.02	0.08	0.03
eigenvector 17	0.02	-0.02	0.02	-0.06	-0.23	0.01	0.07	0.04	0.05	0.00	-0.76	-0.04	0.60	0.01	0.03	-0.01	-0.01

Cuadro A.13: Eigenvectores para el indicador de riqueza, 2012. Oruro

	mp2	rev1	mt1	mps1	mps3	mps6	sal	pal	pa5	ssl	ss2	td1	ee1	cec1	cc2	cc5	tnv1	act2	act3	act5	er2	er5
eigenvector 1	-0.13	0.09	0.20	-0.29	0.07	0.11	0.26	0.30	-0.16	0.17	0.13	0.29	0.26	0.12	0.18	-0.26	-0.14	0.32	0.18	0.28	0.28	-0.17
eigenvector 2	0.03	-0.04	0.26	-0.14	-0.15	0.33	-0.24	-0.02	-0.02	-0.34	0.07	-0.30	0.07	-0.28	0.55	-0.21	-0.11	0.01	-0.27	-0.01	-0.08	0.00
eigenvector 3	-0.25	-0.01	0.15	-0.09	-0.04	0.14	-0.01	-0.05	0.00	0.44	-0.63	-0.25	0.04	0.16	0.15	-0.08	0.37	0.02	0.04	0.03	-0.16	0.07
eigenvector 4	0.60	0.23	0.03	-0.05	-0.04	0.11	-0.17	-0.01	0.03	-0.12	0.06	-0.09	0.22	0.38	-0.10	0.12	0.23	0.23	-0.00	0.36	-0.20	0.15
eigenvector 5	0.39	-0.00	-0.07	0.18	0.16	-0.58	0.06	0.06	-0.04	0.20	-0.28	-0.08	0.09	-0.29	0.31	-0.13	-0.27	0.10	-0.02	0.14	-0.09	-0.00
eigenvector 6	0.58	-0.01	0.17	-0.06	-0.08	0.31	0.21	0.18	-0.09	0.24	-0.14	0.14	-0.12	-0.13	-0.08	0.01	0.09	-0.21	-0.19	-0.38	0.21	-0.18
eigenvector 7	-0.00	0.18	-0.08	-0.16	-0.06	0.38	-0.16	-0.15	0.14	0.26	-0.16	0.10	0.03	-0.41	-0.28	0.18	-0.48	0.07	0.09	0.23	-0.11	0.15
eigenvector 8	0.10	-0.18	-0.08	0.09	-0.02	-0.01	-0.46	-0.29	0.25	0.06	-0.07	0.01	-0.05	-0.05	0.01	-0.08	0.17	0.14	0.12	0.17	0.57	-0.39
eigenvector 9	-0.02	-0.15	0.27	0.04	0.01	-0.07	0.13	0.12	-0.10	-0.14	0.08	-0.03	0.09	-0.63	-0.25	0.07	0.51	0.08	0.20	0.23	-0.08	0.00
eigenvector 10	0.03	0.26	0.66	-0.18	0.26	-0.24	-0.19	-0.26	0.24	0.06	0.11	0.19	-0.12	0.04	-0.00	-0.07	-0.03	-0.14	0.19	-0.19	-0.08	0.09
eigenvector 11	0.15	0.10	-0.37	-0.02	-0.05	0.18	0.27	-0.23	0.25	0.03	0.15	0.18	-0.21	-0.13	0.36	-0.27	0.24	0.00	0.39	-0.06	-0.09	0.24
eigenvector 12	-0.12	0.76	-0.29	-0.16	0.04	-0.13	-0.24	0.18	-0.16	-0.01	-0.01	-0.05	0.06	-0.20	0.01	-0.04	0.23	-0.10	-0.09	-0.13	0.12	-0.12
eigenvector 13	0.01	-0.24	-0.04	0.08	-0.05	-0.02	-0.52	0.25	-0.31	0.22	0.10	0.33	-0.01	-0.02	0.08	-0.08	0.09	-0.02	-0.00	-0.07	0.08	0.55
eigenvector 14	-0.11	0.06	0.03	0.02	-0.01	-0.09	0.23	-0.22	0.31	0.13	0.09	0.22	0.13	-0.08	-0.01	-0.01	0.18	0.08	-0.72	0.15	0.17	0.27
eigenvector 15	0.03	0.12	0.12	-0.00	0.02	-0.02	0.21	0.03	-0.02	-0.20	-0.16	-0.39	-0.08	0.02	-0.04	0.09	-0.11	-0.02	0.18	0.01	0.61	0.53
eigenvector 16	-0.09	0.31	0.26	0.61	-0.60	-0.00	0.04	0.07	0.04	0.09	0.05	0.11	-0.04	0.04	0.12	0.07	-0.06	-0.04	0.10	0.11	0.01	-0.06
eigenvector 17	-0.01	0.10	0.06	0.08	0.18	0.09	0.06	-0.28	-0.53	0.01	0.01	0.04	-0.54	0.04	-0.01	-0.14	0.00	-0.15	-0.15	0.45	-0.00	-0.06
eigenvector 18	-0.00	-0.09	-0.03	-0.21	-0.00	-0.05	-0.03	0.41	0.35	0.02	0.01	0.02	-0.13	0.04	0.18	0.18	0.02	-0.61	0.01	0.44	0.02	-0.03
eigenvector 19	-0.05	0.07	0.05	0.14	0.20	0.08	-0.08	0.45	0.31	-0.04	-0.05	-0.04	-0.57	-0.01	-0.03	-0.00	0.01	0.50	-0.12	-0.04	-0.09	-0.00
eigenvector 20	-0.05	0.08	0.01	0.54	0.64	0.36	-0.01	0.04	0.01	0.01	0.03	0.31	-0.01	0.14	0.09	0.02	-0.17	0.03	-0.00	0.03	-0.01	0.01
eigenvector 21	0.03	0.01	-0.01	0.13	-0.04	0.04	-0.04	0.14	0.16	-0.10	-0.08	-0.09	0.10	0.04	-0.43	-0.80	-0.08	-0.21	-0.04	0.07	-0.05	0.05
eigenvector 22	-0.02	0.02	-0.00	-0.02	-0.02	-0.00	-0.01	-0.01	-0.01	-0.57	-0.59	0.56	0.03	0.02	0.04	0.04	0.01	0.00	0.01	0.01	-0.00	0.02

Cuadro A.14: Eigenvectores para el indicador de riqueza, 2012. Potosi

	rev1	mt1	mt4	mps1	mps6	pa1	pa6	ss1	ss2	td1	ee1	cc2	tnv1	act2	act3	act5	act6	er2	er4
eigenvector 1	0.18	0.20	-0.24	-0.33	0.24	0.26	-0.12	0.12	0.17	0.29	0.28	0.26	-0.14	0.33	0.13	0.31	0.09	0.27	-0.13
eigenvector 2	-0.09	-0.62	0.42	0.12	-0.24	0.19	0.04	0.14	0.13	0.30	0.00	-0.15	-0.07	0.11	0.16	0.15	0.04	0.23	-0.22
eigenvector 3	0.02	-0.21	0.11	-0.21	0.36	-0.14	0.10	-0.50	0.49	0.03	-0.17	0.15	-0.30	-0.17	-0.14	-0.18	-0.12	0.07	-0.08
eigenvector 4	0.29	-0.36	0.21	-0.34	0.56	0.00	0.07	0.21	-0.28	-0.15	0.05	-0.12	0.28	0.02	-0.04	0.00	0.09	-0.23	0.03
eigenvector 5	-0.09	0.14	-0.13	-0.14	0.13	0.24	-0.05	0.36	-0.12	0.26	-0.40	-0.09	0.01	-0.39	0.04	-0.42	-0.05	0.24	-0.29
eigenvector 6	-0.00	-0.03	-0.00	-0.11	-0.07	-0.69	0.31	0.23	-0.29	0.01	-0.12	0.28	-0.29	0.04	0.10	0.11	0.05	0.18	-0.19
eigenvector 7	0.15	0.21	-0.12	-0.03	0.03	-0.29	0.22	-0.16	0.28	0.14	-0.05	-0.63	0.29	0.04	0.22	0.11	0.24	0.05	-0.22
eigenvector 8	0.80	-0.01	-0.02	0.04	-0.28	0.04	-0.01	0.06	-0.01	-0.01	0.11	-0.17	-0.37	-0.09	-0.01	-0.18	-0.22	-0.09	-0.04
eigenvector 9	0.06	0.03	0.01	0.15	-0.05	0.09	-0.02	-0.17	-0.07	-0.30	0.13	0.20	0.23	0.05	-0.22	0.01	-0.04	-0.03	-0.82
eigenvector 10	-0.36	0.02	-0.09	-0.10	0.16	-0.07	-0.01	0.11	-0.09	-0.01	0.29	-0.46	-0.29	0.10	-0.17	0.10	-0.59	-0.07	-0.14
eigenvector 11	-0.18	-0.03	-0.06	-0.13	0.00	0.15	-0.22	-0.09	-0.12	-0.25	-0.05	-0.20	-0.54	-0.04	0.21	0.03	0.54	-0.31	-0.17
eigenvector 12	-0.09	-0.08	-0.04	0.00	-0.05	-0.20	-0.06	0.14	0.10	0.15	0.52	-0.06	-0.00	0.05	-0.41	-0.51	0.40	0.13	0.04
eigenvector 13	0.11	-0.09	-0.02	-0.05	0.01	-0.11	-0.46	-0.20	-0.23	-0.25	-0.09	-0.21	0.01	-0.13	-0.22	0.23	0.05	0.64	0.11
eigenvector 14	-0.05	-0.13	-0.03	-0.15	-0.06	-0.25	-0.46	-0.09	0.03	-0.08	0.21	0.10	0.18	0.04	0.63	-0.34	-0.23	-0.01	-0.05
eigenvector 15	0.03	-0.15	-0.18	-0.02	-0.07	-0.26	-0.41	0.20	0.21	0.29	-0.02	0.09	0.10	-0.42	-0.21	0.38	0.02	-0.37	-0.09
eigenvector 16	-0.06	-0.08	-0.11	-0.07	-0.04	0.16	0.37	-0.10	-0.06	-0.12	0.48	0.06	0.04	-0.66	0.23	0.16	0.03	0.15	0.05
eigenvector 17	0.04	0.46	0.68	0.29	0.29	-0.10	-0.18	0.11	0.04	0.06	0.18	0.02	-0.12	-0.20	0.10	0.07	0.04	0.02	-0.03
eigenvector 18	0.08	-0.24	-0.40	0.71	0.46	-0.04	-0.01	0.05	-0.02	0.02	0.01	0.01	-0.11	0.03	0.17	-0.03	0.03	0.06	0.00
eigenvector 19	0.03	0.01	0.03	0.02	0.03	0.00	-0.03	-0.53	-0.56	0.61	0.06	0.00	0.01	-0.00	-0.02	-0.03	0.01	-0.13	-0.04

Cuadro A.15: Eigenvectores para el indicador de riqueza, 2012. Tarija

	mp1	rev1	mps5	sa1	pa1	td1	td3	tnv1	act2	act3	act5	er2	er5
eigenvector 1	0.29	0.13	0.18	0.33	0.20	0.45	-0.20	-0.16	0.28	0.23	0.28	0.39	-0.29
eigenvector 2	0.04	0.17	0.08	0.11	0.10	-0.25	0.41	0.43	0.42	0.22	0.46	-0.20	0.20
eigenvector 3	0.07	-0.00	-0.16	-0.41	-0.08	-0.15	0.26	-0.68	0.28	-0.20	0.34	0.10	0.01
eigenvector 4	0.41	-0.17	0.11	-0.54	-0.37	-0.07	-0.03	0.42	0.00	0.09	0.01	0.36	-0.22
eigenvector 5	0.62	0.14	0.47	0.16	-0.06	-0.15	0.25	-0.27	-0.23	0.07	-0.25	-0.21	0.16
eigenvector 6	0.21	-0.03	-0.31	0.36	0.15	-0.14	0.41	0.17	-0.04	-0.55	-0.13	0.36	-0.20
eigenvector 7	0.46	-0.13	-0.24	0.02	-0.06	0.23	-0.40	0.12	0.20	-0.42	0.16	-0.42	0.24
eigenvector 8	0.10	-0.65	-0.31	0.38	-0.32	-0.14	0.06	-0.15	0.01	0.42	0.05	-0.01	0.02
eigenvector 9	0.28	0.39	-0.66	-0.19	0.23	0.08	0.06	-0.01	-0.15	0.43	-0.17	-0.03	0.05
eigenvector 10	0.09	-0.54	0.13	-0.27	0.76	0.03	0.07	0.02	-0.02	0.06	-0.05	0.00	0.10
eigenvector 11	-0.03	0.03	0.01	0.04	-0.06	0.05	-0.08	0.01	0.00	-0.01	-0.04	0.55	0.82
eigenvector 12	0.04	0.02	-0.02	0.04	0.07	-0.10	-0.09	0.03	-0.72	-0.04	0.67	0.04	0.00
eigenvector 13	0.09	0.12	-0.01	0.10	0.20	-0.76	-0.55	-0.06	0.16	0.02	-0.05	0.12	-0.08

Cuadro A.16: Eigenvectores para el indicador de riqueza, 2012. Santa Cruz

	mp1	rev1	mt2	mps5	sa1	pa1	td1	td3	cc2	act2	act3	act5	act6	er2	er5
eigenvector 1	0.26	0.25	0.23	0.26	0.31	0.21	0.31	-0.21	0.22	0.26	0.25	0.26	0.18	0.33	-0.29
eigenvector 2	-0.14	-0.07	-0.26	0.13	0.14	-0.15	0.48	-0.60	-0.22	-0.32	0.09	-0.31	0.00	-0.06	-0.00
eigenvector 3	-0.02	0.07	-0.02	0.38	0.09	-0.13	-0.10	0.12	-0.05	0.05	0.40	0.09	0.51	-0.50	0.35
eigenvector 4	-0.20	-0.28	-0.57	-0.33	-0.20	0.04	0.09	-0.10	0.01	0.34	0.23	0.41	0.21	0.11	-0.06
eigenvector 5	-0.00	-0.01	0.58	-0.23	-0.50	-0.01	0.18	-0.41	-0.06	0.16	0.01	0.20	-0.01	-0.24	0.18
eigenvector 6	-0.10	-0.04	0.14	0.19	-0.53	-0.14	-0.14	0.10	-0.14	-0.23	0.22	-0.22	0.38	0.47	-0.26
eigenvector 7	0.17	0.26	-0.36	0.56	-0.49	-0.01	0.03	-0.06	0.09	0.11	-0.12	0.15	-0.38	-0.06	0.06
eigenvector 8	0.33	0.33	-0.22	-0.23	-0.12	0.02	0.02	-0.13	0.20	-0.04	-0.52	-0.14	0.56	-0.07	0.01
eigenvector 9	0.22	0.54	-0.13	-0.44	-0.11	-0.02	0.09	0.16	-0.02	-0.14	0.55	-0.16	-0.20	-0.06	0.01
eigenvector 10	0.39	-0.46	-0.05	-0.00	-0.08	0.16	-0.28	-0.26	0.54	-0.12	0.28	-0.26	-0.06	-0.09	0.02
eigenvector 11	-0.12	-0.17	0.02	0.06	-0.19	0.53	0.59	0.44	0.17	-0.06	-0.03	-0.20	0.07	-0.14	0.05
eigenvector 12	0.20	-0.20	0.06	-0.03	0.00	-0.76	0.40	0.28	0.32	0.04	-0.04	0.05	-0.04	0.05	0.02
eigenvector 13	0.68	-0.30	-0.03	-0.00	-0.00	0.10	0.09	0.12	-0.61	-0.11	-0.02	0.17	0.02	0.01	0.02
eigenvector 14	0.01	0.04	-0.01	-0.01	0.05	0.08	0.03	-0.04	0.06	-0.06	0.01	0.02	0.01	0.55	0.82
eigenvector 15	-0.08	0.04	0.01	-0.02	-0.00	0.06	0.00	0.01	0.20	-0.76	-0.02	0.60	0.02	-0.05	-0.06

Cuadro A.17: Eigenvectores para el indicador de riqueza, 2012. Beni

	mpl	mp5	rev1	mps1	mps5	sa1	pa1	pa5	td2	td3	ee1	cc2	act2	act3	act5	act8	er2	er5
eigenvector 1	0.32	-0.23	0.28	-0.31	0.18	0.25	0.23	-0.12	0.15	-0.20	0.25	0.28	0.26	0.15	0.26	0.16	0.26	-0.25
eigenvector 2	0.24	-0.30	0.28	-0.24	0.16	0.13	0.03	-0.16	0.10	-0.30	-0.27	-0.23	-0.39	0.00	-0.35	-0.30	-0.19	0.15
eigenvector 3	-0.23	0.28	-0.27	0.21	-0.11	0.28	0.53	-0.43	0.06	-0.27	0.03	0.08	-0.08	-0.08	-0.09	-0.25	0.11	-0.12
eigenvector 4	0.08	-0.02	-0.06	-0.01	-0.02	-0.14	-0.22	0.24	-0.02	0.00	0.06	0.09	-0.20	-0.13	-0.28	-0.36	0.56	-0.52
eigenvector 5	-0.11	0.19	-0.15	0.05	0.02	-0.14	-0.31	0.07	0.46	-0.65	-0.15	-0.02	-0.02	0.12	-0.01	0.36	0.06	-0.05
eigenvector 6	-0.08	0.23	0.06	-0.12	0.21	0.25	0.10	-0.04	-0.31	0.20	-0.25	-0.11	-0.24	0.17	-0.31	0.58	0.19	-0.19
eigenvector 7	-0.33	0.49	0.23	-0.26	0.35	0.28	-0.18	0.23	-0.01	-0.04	0.07	-0.04	0.08	0.21	0.15	-0.40	-0.07	0.04
eigenvector 8	-0.09	-0.21	0.10	0.25	-0.03	0.04	0.10	0.07	-0.13	-0.13	-0.06	-0.75	0.08	0.14	0.38	-0.04	0.26	-0.14
eigenvector 9	0.04	-0.18	-0.06	0.28	-0.15	0.70	-0.09	0.47	0.04	-0.10	0.13	0.03	0.07	-0.20	-0.19	0.09	-0.14	0.01
eigenvector 10	-0.15	0.18	0.37	-0.21	-0.12	-0.23	0.44	0.31	0.26	0.00	0.13	-0.20	0.04	-0.48	-0.15	0.15	0.05	0.07
eigenvector 11	-0.38	-0.15	0.66	0.48	-0.03	-0.05	-0.15	-0.22	0.05	0.02	0.01	0.25	0.00	0.09	-0.14	0.03	0.03	-0.05
eigenvector 12	-0.01	-0.09	-0.02	0.14	0.08	-0.30	0.41	0.46	-0.34	-0.31	0.06	0.22	-0.04	0.43	-0.13	-0.04	-0.15	0.02
eigenvector 13	0.07	-0.01	-0.09	0.14	0.01	0.05	0.19	0.13	0.63	0.43	0.04	-0.07	-0.14	0.49	-0.11	-0.06	0.13	0.14
eigenvector 14	0.08	-0.02	-0.14	0.18	0.53	-0.13	-0.04	-0.16	0.02	0.00	0.27	-0.25	0.50	-0.06	-0.47	-0.01	-0.06	-0.02
eigenvector 15	-0.08	0.02	0.06	-0.33	-0.64	0.03	-0.09	-0.10	-0.07	-0.04	0.04	-0.15	0.40	0.35	-0.36	-0.06	-0.06	-0.10
eigenvector 16	-0.02	-0.00	0.00	0.00	0.02	-0.06	0.05	-0.01	0.16	0.11	0.09	-0.10	-0.15	0.02	0.10	0.03	-0.62	-0.72
eigenvector 17	0.56	0.48	0.24	0.25	-0.18	-0.02	-0.11	-0.09	-0.12	-0.09	0.41	-0.14	-0.23	0.08	0.01	0.04	-0.03	0.08
eigenvector 18	-0.39	-0.27	-0.15	-0.22	0.02	-0.00	-0.11	-0.10	-0.06	-0.06	0.68	-0.10	-0.39	0.06	-0.03	0.13	0.04	0.14

Cuadro A.18: Eigenvectores para el indicador de riqueza, 2012. Pando

	mp5	rev1	mt1	mps1	sa2	pa1	pa5	td2	td3	ee1	cc2	tnv2	act2	act3	act5	er2	er5
vector propio 1	-0.14	0.18	0.23	-0.24	0.16	0.33	-0.14	0.12	-0.15	0.35	0.33	0.10	0.31	0.17	0.33	0.29	-0.29
eigenvector 2	-0.56	0.38	-0.22	0.09	-0.28	-0.04	-0.04	0.20	-0.52	-0.12	-0.13	0.05	-0.17	0.13	-0.11	0.03	-0.05
eigenvector 3	-0.11	0.11	0.00	-0.03	-0.50	-0.38	0.44	-0.01	0.18	0.15	0.14	-0.11	0.38	0.15	0.31	-0.12	0.15
eigenvector 4	0.14	-0.01	0.61	-0.43	-0.27	-0.20	0.15	0.07	-0.17	-0.17	0.17	0.00	-0.30	-0.10	-0.26	0.09	-0.16
eigenvector 5	0.56	-0.34	-0.23	0.02	-0.15	-0.03	0.01	0.26	-0.61	-0.00	-0.02	0.03	0.15	0.11	0.17	-0.02	0.02
eigenvector 6	-0.06	0.13	0.22	-0.33	0.16	0.08	-0.27	0.12	-0.13	-0.02	0.02	-0.04	0.06	0.07	0.08	-0.52	0.63
eigenvector 7	0.15	0.06	-0.32	-0.49	-0.35	0.00	-0.47	-0.13	0.29	-0.21	-0.13	0.02	-0.01	0.30	0.05	0.08	-0.14
eigenvector 8	-0.04	-0.07	-0.16	0.11	0.07	0.01	-0.10	0.14	0.04	-0.21	0.63	-0.63	-0.07	0.21	-0.15	0.06	0.04
eigenvector 9	-0.02	-0.12	0.52	0.47	-0.23	-0.05	-0.45	-0.02	-0.00	-0.01	-0.28	-0.26	0.13	0.20	0.10	-0.06	-0.13
eigenvector 10	0.07	0.07	0.11	-0.07	0.35	0.07	0.42	0.02	0.02	-0.19	-0.34	-0.12	-0.13	0.69	0.07	0.02	-0.05
eigenvector 11	-0.12	-0.01	-0.09	-0.31	0.28	-0.29	-0.01	0.21	-0.05	0.02	-0.33	-0.47	0.27	-0.32	0.01	-0.15	-0.37
eigenvector 12	-0.10	-0.07	0.05	-0.09	0.09	-0.18	-0.06	-0.11	-0.09	-0.17	-0.17	-0.16	-0.01	-0.16	0.29	0.70	0.47
eigenvector 13	-0.14	-0.11	0.01	0.07	0.25	-0.36	-0.09	-0.22	-0.09	-0.50	0.27	0.31	-0.02	0.00	0.43	-0.20	-0.23
eigenvector 14	0.04	0.04	-0.00	0.07	-0.06	0.03	-0.01	0.66	0.31	0.03	-0.03	0.02	-0.47	-0.14	0.46	-0.02	-0.03
eigenvector 15	-0.01	-0.04	-0.06	-0.09	-0.10	0.15	0.06	-0.52	-0.22	0.33	-0.01	-0.35	-0.47	-0.08	0.36	-0.18	-0.08
eigenvector 16	-0.02	-0.09	-0.06	-0.01	0.22	-0.62	-0.23	0.04	-0.00	0.55	0.07	0.14	-0.23	0.29	-0.16	0.07	0.05
eigenvector 17	0.50	0.79	0.01	0.19	0.10	-0.18	-0.08	-0.10	-0.05	-0.04	0.04	-0.08	-0.00	-0.11	0.04	0.03	-0.02

Bibliografía

- Abeyasekera, S. (1995). Multivariate methods for index construction. *Household Surveys in Developing and Transition Countries: Design, Implementation and Analysis*, 1–21. Retrieved from <http://unstats.un.org/unsd/hhsurveys/> doi: 10.1016/j.econ.2011.08.014
- Cattell, R. B. (1966). The scree test for the number of factors. *University of Illinois*, 245-276.
- Córdova, a. (2009). Methodological Note: Measuring Relative Wealth using Household Asset Indicators. *Insights Series No. I0806*, 1–9. Retrieved from <http://vanderbilt.edu/lapop/insights/I0806en.pdf>
- Crittenden, J. C. (2012). Index. , 1869–1901. doi: 10.1002/9781118131473.index
- de Leeuw, J. (2003). Principal component analysis of binary data: Applications to roll-call analysis. *Mathematics Subject Classification*, 32. Retrieved from <http://escholarship.org/uc/item/7n7320n0.pdf> <http://escholarship.org.ezproxy.lib.indiana.edu/uc/item/7n7320n0>
- de Leeuw, J. (2006). Principal component analysis of binary data by iterated singular value decomposition. *Computational Statistics & Data Analysis*, 50(1), 21–39. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0167947304002300> doi: 10.1016/j.csda.2004.07.010
- Dunteman, G. (1989). *Principal components analysis* (No. n.^o 69). SAGE Publications. Retrieved from <https://books.google.com.bo/books?id=Pzwt-CMMt4UC>
- Feres, J. C., & Mancero, X. (2001). El Método de las Necesidades Básicas Insatisfechas (NBI) y sus Aplicaciones en América Latina. *Estudios Estadísticos y Prospectivos*, 7, 61–100.
- Filmer, D., Pritchett, L., & Resources, W. (1998). *Estimating wealth effects without expenditure data, or tears: An application to educational enrollments in states of india* (No. v. 1994). World Bank, Development Research Group, Poverty and Human Resources.
- Filmer, D., & Scott, K. (2008). Assessing Asset Indices (WPS 4605). (April), –.
- Freeman, J., & Jackson, J. E. (1992). *A User's Guide to Principal Components* (Vol. 43) (No. 6). doi: 10.2307/2583020
- Giorgio Alleva, A. G. e. (2016). *Topics in theoretical and applied statistics* (1st ed.). Springer International Publishing. Retrieved from <http://gen.lib.rus.ec/book/index.php?md5=8D271CAB00B0B8ADEB97EE6DB96EF653>
- Gomez, A., Alvarez, G., Lucarini, A., & Olmos, F. (1994). Las necesidades básicas insatisfechas: Sus deficiencias técnicas y su impacto en la definición de políticas sociales. *INDEC - Argentina*, 128.
- Gower, J. C. (1966). *Some distance properties of latent root and vector methods used in multivariate analysis*. *Biometrika*.

- Gragmolati, M., & Watts, C. (2008). Annex 5 : Principal Components Analysis and the Wellbeing Index. (2000).
- Harttgen, K., & Vollmer, S. (2011). Inequality Decomposition without Income or Expenditure Data: Using an Asset Index to Simulate Household Income. *Human Development*.
- Instituto Nacional de Estadística - Bolivia. (2005). *CÁLCULO DEL INDICADOR DE NECESIDADES BÁSICAS INSATISFECHAS EN BOLIVIA 1992 Y 2001* (Tech. Rep.).
- Johnson, R., & Wichern, D. (2007). *Applied multivariate statistical analysis*. Pearson Prentice Hall. Retrieved from <https://books.google.com.bo/books?id=gFWcQgAACAAJ>
- Jolliffe, I. (2013). *Principal component analysis*. Springer New York.
- Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*, 30(3), 487. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa501/full> doi: 10.2307/1270093
- Jolliffe, I. T., Trendafilov, N. T., & Uddin, M. (2003). A Modified Principal Component Technique Based on the LASSO. *Journal of Computational and Graphical Statistics*, 12(3), 531–547. Retrieved from <http://www.tandfonline.com/doi/abs/10.1198/1061860032148> doi: 10.1198/1061860032148
- Kolenikov, S., & Angeles, G. (2004). The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices. *Chapel Hill: Carolina Population Center, University of North Carolina.*, 1–59.
- Lanjouw, P., Elbers, C., & Lanjouw, J. (2000). Micro-Level estimations of Welfare. doi: 10.1596/1813-9450-2911
- Lee, S., Huang, J. Z., & Hu, J. (2010). Sparse logistic principal components analysis for binary data. *Annals of Applied Statistics*, 4(3), 1579–1601. doi: 10.1214/10-AOAS327
- McKenzie, D. (2004). Measuring Inequality with Aset Indicators. *BREAD working paper No. 042*.
- Ministerio de Autonomías. (2013). Agenda Patriótica 2025 ¿Quién hace qué? Serie Autonomías para la Gente. *Nro 6*.
- Ministerio de Planificación del Desarrollo. (2015). *Plan de desarrollo económico y social, En el marco del desarrollo integral para vivir bien, 2016-2020*.
- Montgomery, M. R., Gragnolati, M., Burke, K. A., & Paredes, E. (2000). Measuring living standards with proxy variables. *Demography*, 37(2), 155-174.
- Murtagh, F. (2007). *Multiple correspondence analysis and related methods* (Vol. 72) (No. 2). doi: 10.1007/s11336-006-1579-x
- Niitsuma, H., & Okada, T. (2005). Covariance and PCA for Categorical Variables. *Advances in Knowledge Discovery and Data Mining*, 3518, 523–528. Retrieved from http://dx.doi.org/10.1007/11430919_{_}61 doi: 10.1007/11430919_61
- Rutstein, S., Johnson, K., & (Programme), O. M. M. (2004). *The dhs wealth index*. ORC Macro, MEASURE DHS.
- Saisana, M., Tarantola, S., for the Protection, I., of the Citizen. Technological, S., & Unit, E. R. M. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development*. European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit.
- Sampieri, R., Collado, C., Lucio, P., & de la Luz rev. téc Casas Pérez, M. (1991).

- Metodología de la investigación*. Universidad de Cuenca. Retrieved from <https://books.google.com.bo/books?id=hk0md0gghywc>
- Schein, A., Saul, L., & Ungar, L. (2003). A generalized linear model for principal component analysis of binary data. *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics*, 546431. Retrieved from <http://www.andrewschein.com/publications/aistat2003talk.pdf>
- Shen, H., & Huang, J. Z. (2008). Sparse principal component analysis via low rank matrix approximation. *Journal of Multivariate Analysis*, 99(6), 1015–1034.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321–327. doi: 10.1007/BF02293557
- Wagstaff, A., O'Donnell, O., Wagstaff, A., & Lindelow, M. (2008). Measurement of Living Standards. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and their Implementation*, 69–82. Retrieved from <http://siteresources.worldbank.org/INTPAH/Resources/Publications/459843-1195594469249/HealthEquityCh6.pdf>