

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA



**ESTRUCTURA QUE POSEE LA TASA DE
DESOCUPACIÓN EN LA REGIÓN DEL BÍO BÍO**

Proyecto de título
GEOVANNI HUMBERTO MORENO VILLA

Profesor Guía: Arturo Mora Cerna
Profesor colaborador: Sebastián Niklitschek Soto
Profesora consejera: Katia Sáez Carrillo

Concepción, Chile
2020

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA



Proyecto de título

**ESTRUCTURA QUE POSEE LA TASA DE
DESOCUPACIÓN EN LA REGIÓN DEL BÍO BÍO**

Profesor Guía	: Arturo Mora Cerna	Firma
Profesor Colaborador	: Sebastián Niklitschek Soto	Firma
Profesora Consejera	: Katia Sáez Carrillo	Firma
Nombre Memorante	: Geovanni Moreno Villa	Firma

Índice general

1. Introducción.	1
2. Marco Teórico.	2
2.1. Nueva Encuesta Nacional de Empleo.	2
2.1.1. El desempleo (Definición y Relevancia).	2
2.1.2. Importancia de la tasa de desempleo a nivel país y región.	3
2.1.3. La nueva encuesta (Mejoras).	3
2.1.4. Diseño muestral.	4
2.1.5. Medición de la Encuesta.	5
2.1.6. Definiciones o Conceptos.	5
2.1.7. Principales indicadores de razón.	8
2.1.8. Factor de expansión.	8
2.2. Análisis de Correspondencias.	10
2.2.1. Variables categóricas.	10
2.2.2. Análisis de Correspondencia Simple (ACS).	13
2.2.2.1. Perfiles filas y perfiles columnas.	13
2.2.2.2. Distancias Chi-Cuadrado (χ^2) e Inercia.	15
2.2.2.3. Las dos nubes de perfiles y sus ajustes.	16
2.2.2.4. Relaciones de transición.	18
2.2.2.5. Contribuciones y representación.	19
2.2.2.6. Interpretaciones.	20
2.2.3. Análisis de Correspondencias Múltiples (ACM).	21
2.2.3.1. Tipos de tablas.	21
2.2.3.2. Perfiles, Distancias, Inercia, Relaciones de Transición.	24
2.2.3.3. Contribuciones, Medidas de Discriminación, Variables en ACM.	26
2.2.3.4. Resultados de ACM para una matriz de Burt, comparación y similitudes con TDC.	29
2.2.3.5. Otras características, propiedades, conclusiones e interpretaciones.	31
3. Objetivos.	33
3.1. Objetivo General:	33
3.2. Objetivos Específicos:	33
4. Materiales y Métodos.	34

<i>ÍNDICE GENERAL</i>	III
5. Resultados.	39
5.1. Análisis regional descriptivo.	39
5.2. Análisis regional mediante Tablas de Contingencias.	42
5.3. Análisis regional mediante ACS	49
5.4. Análisis regional mediante ACM.	72
5.5. Análisis provincial mediante ACM.	92
6. Conclusión, discusión y proyecciones.	97
7. Anexo.	104
7.1. ACM Provincia de Concepción.	104
7.2. ACM Provincia de Arauco.	108
7.3. ACM Provincia del Biobío.	112



Capítulo 1

Introducción.

La nueva encuesta nacional de empleo (NENE) es el nuevo instrumento de medición de los indicadores coyunturales del empleo y desempleo. Aporta información sobre características de la población económicamente activa (PEA) o fuerza de trabajo, y sobre la población inactiva o fuera de la fuerza del trabajo a nivel nacional, regional, provincial, por género, por grupos etáreos, entre otros. Esta encuesta sigue las recomendaciones adoptadas por las Conferencias Internacionales de Estadística del Trabajo de la OIT. Posee una muestra trimestral de aproximadamente 36.000 hogares en todo el país.

Desde la implementación de la NENE (trimestre móvil Enero-Febrero-Marzo 2010), la Región del Biobío históricamente ha tenido índices de desocupación mayores que el nacional (en condiciones normales), tanto en el promedio anual como por trimestre móvil, excepto algunos meses de los últimos años [1]. Incluso, cuando Ñuble era parte de la Región del Biobío (registros hasta Sept-Oct-Nov 2018), e incluso con Factor de Expansión y proyecciones de población en base al Censo 2002 (registros hasta Oct-Nov-Dic 2019). Actualmente, el índice de desocupación de 2019 en el país fue de 7,2% y en la Región del Biobío fue de 7,3%, en base a proyecciones de población del Censo 2017 y mejoras en el Factor de Expansión. De manera contraria, la tasa de ocupación de la región se encuentra bajo el promedio nacional, siendo por género las tasas más bajas de las regiones. De esta forma la Región del Biobío posee niveles críticos de empleo y desempleo.

Surge entonces el interés por modelar la tasa de empleabilidad de la Región del Biobío y de sus 3 provincias, con el objeto de determinar aquellos factores que están incidiendo en la empleabilidad local, lo que permitiría realizar acciones para controlar de manera favorable dichos factores.

En este contexto, aparece como interesante determinar alguna estructura en la tasa de desocupación de la región y sus provincias, para identificar grupos con características comunes, lo que facilitaría la generación de políticas públicas para mejorar los índices de desocupación de la región y sus provincias. Así, una vez determinados los factores o la estructura en la tasa de desocupación, los resultados pueden ser entregados a la entidad respectiva para que trate de cubrir las necesidades coyunturales que provocan los altos índices de desempleo, y los bajos índices de ocupación de nuestra región y sus provincias.

Capítulo 2

Marco Teórico.

2.1. Nueva Encuesta Nacional de Empleo.

2.1.1. El desempleo (Definición y Relevancia).

El desempleo, también llamado desocupación, cesantía o paro, es la situación en la que se encuentran las personas o ciudadanos que teniendo la edad, capacidad y deseo de trabajar carecen de empleo. Otros lo definen como la incapacidad de una economía para absorber a toda la fuerza laboral. Es un fenómeno involuntario, ya que las personas no deciden ser contratadas por las empresas. La situación opuesta al desempleo es el empleo. Aparte de la población desempleada y su opuesto, población empleada o población que se encuentra trabajando, los países cuentan con una población inactiva, que son las personas que no están disponibles para trabajar, ya sea por estudios (niños), por jubilados, enfermedades u otra causa legalmente establecida.

Para medir el desempleo de un país se utiliza la tasa de desempleo, que es una proporción de personas que buscan trabajo sin encontrarlo, con respecto al total de personas que forman la fuerza laboral. De manera análoga, la tasa de desempleo no es el único indicador del mercado laboral, otro muy utilizado es la tasa de ocupación, que es una proporción de trabajadores empleados con respecto a la población total en edad de trabajar.

La sociedad entiende y está de acuerdo que el desempleo es un grave problema que causa pérdidas de todo tipo, entre las cuales es la pérdida de producción e ingreso, pérdida de capital humano (costes psicológicos u otro que afecte la salud del trabajador y de paso al entorno familiar), o pérdida del sector público (habrá menos personas que paguen impuestos, por ende empeora la calidad de los servicios públicos de las ciudades). En ciertos episodios de la historia, países han experimentado altas tasas de desempleo y por periodos extensos, llegando a crisis económicas a nivel nación.

2.1.2. Importancia de la tasa de desempleo a nivel país y región.

El indicador de desempleo es muy importante, ya que con él se empieza a ordenar la realidad de la coyuntura nacional, regional, por sectores rurales y urbanos, etc.

Chile, resulta ser un país con un alto dinamismo en su fuerza laboral (PEA), así se demuestra en las tasas de desempleo/empleo de los últimos años, esto debido a procesos y factores naturales de la economía. En particular, se ha caracterizado una baja del empleo asalariado y un alza en el empleo por cuenta propia.

La Región del Biobío no está exenta del dinamismo de la fuerza de trabajo, el desempeño económico es lento, lo que se traduce en tasas de desempleos mayores al promedio nacional en los últimos años, lo que a su vez se refleja en la fuga de profesionales al norte y centro del país principalmente por los bajos sueldos y la baja movilidad laboral. En la región existe una alta tasa de matrículas a educación superior, pero con baja retención, ya que no hay incentivos que capturen a los profesionales al estar egresados y aptos para buscar empleos.

Otro dato relevante, es que el PIB de la región para el año 2015 fue de un 7,7%. Hace décadas la región aportaba alrededor del 12% a 14% al PIB, lo que habla de un empobrecimiento de la región.

Es por estos motivos que el interés de la tasa de desempleo se centrará en nuestra región.

Por todo el dinamismo de la fuerza de trabajo existente en el país y en cada región, especialmente en la Región del Biobío, se requiere de instrumentos modernos para realizar un estudio e identificar las condiciones laborales existentes y tasas de ocupados, desocupados o fuera de la fuerza de trabajo. Es por esto, que ciertos organismos internacionales recomiendan y exigen aplicar nuevas y modernas metodologías de estudio, tal que representen de la mejor forma la evolución de la fuerza laboral.

2.1.3. La nueva encuesta (Mejoras).

La Nueva Encuesta Nacional de Empleo (desde ahora “NENE”), es el nuevo instrumento de medición de la tasa empleo y desempleo del país. Se comenzó a aplicar por primera vez a partir del trimestre móvil enero-marzo 2010 y reemplazó a la antigua Encuesta Nacional de Empleo (ENE) que estaba vigente desde 1996. Implementó una serie de modificaciones que aportan nueva información sobre características del empleo y sobre las personas inactivas, o sea, estas modificaciones están destinadas principalmente a definir con mayor precisión cada categoría de la situación ocupacional de las personas. Algunas modificaciones fueron los cambios conceptuales en la clasificación de la condición laboral, cambios en la forma de medición, incorporación de nuevos indicadores, entre otros.

Las nuevas modificaciones se basan en conceptos y criterios acorde a los estándares internacionales y las mejores prácticas metodológicas y operativas de los países miembros de la

Organización para la Cooperación y el Desarrollo Económico (OCDE), o sea, la NENE toma en cuenta las actualizaciones emanadas de las resoluciones de las Conferencias Internacionales de Estadísticos del Trabajo (CIET), de la Organización Internacional del Trabajo (OIT).

Algunas de las **principales modificaciones** son:

- a) Referente a la población ocupada, incorporó un registro según rama de actividad, categoría y grupos ocupacionales distinguiendo entre tradicionales y no tradicionales. Incluyendo criterios como el haber trabajado al menos una hora en la semana de referencia y recibir un ingreso por la actividad realizada.
- b) Referente a la población desocupada, incluyó criterios de haber buscado empleo en las últimas 4 semanas y de estar disponible para trabajar en, al menos, las 2 semanas siguientes a la de entrevista.
- c) Clasificación de las personas fuera de la fuerza de trabajo, por razones de inactividad y también por su potencialidad de participar en mercado laboral.

Algunos de los **beneficios** que tendrá la NENE es que permitirá a instituciones públicas y privadas diseñar mejores políticas laborales, esto debido a que proporciona información relevante para el análisis del mercado laboral. Por otra parte, instituciones internacionales podrán realizar estudios de comparabilidad en relación a la fuerza laboral, ya que utilizan metodología similar. Además se podrá conocer el comportamiento de la Población Económicamente Activa (PEA) en los ciclos económicos, con análisis de género, análisis sectorial del empleo, análisis del comportamiento de la PEA e inactivos, estudiar el desempleo juvenil y empleo/desempleo según escolaridad.

El **objetivo** de esta encuesta es clasificar de acuerdo a estándares internacionales definidos por la Organización Internacional del trabajo (OIT), a la población en edad de trabajar con residencia habitual en el país. La encuesta entrega resultados a nivel nacional, según sus niveles de estimación tales como Nacional Urbano-Rural, Regiones Urbano-Rural, provincias y grandes ciudades.

2.1.4. Diseño muestral.

Esta encuesta está basada en una muestra trimestral de aproximadamente 36.000 viviendas en todo el país, seleccionadas por medio de un muestreo probabilístico bietápico, donde en cada vivienda se entrevista a cada uno de los miembros del hogar mayores o iguales a 15 años. El muestreo bietápico tiene como unidad de selección de primera etapa (UPE) las manzanas en el sector urbano y las secciones en el sector rural, y como unidad de segunda etapa (USE) a las viviendas particulares ocupadas, y finalmente como unidad de observación a las personas.

La muestra total obtenida se distribuye en un período trimestral, se divide en tres submuestras de un tamaño aproximadamente similar, cada una se asigna en forma sistemática a uno

de los tres meses del trimestre móvil. De este modo, la vivienda seleccionada tiene cierta cantidad de visitas con cierta periodicidad, dependiendo de si el estrato es de grandes centros urbanos o ciudades (CD), resto de áreas urbanas (RAU) o área rural (R).

En el caso de grandes centros urbanos, las visitas son cada 3 meses hasta completar un total de 6 visitas, a lo largo de 18 meses; el caso de resto de áreas urbanas, las visitas son cada 3 meses hasta completar 9 visitas, a lo largo de 27 meses; en el caso de área rural las visitas son cada 3 meses hasta completar 12 visitas, durante un periodo de 36 meses.

De esta forma se evita el cansancio del informante, y es mecanismo de actualización de la muestra mediante la rotación de las viviendas dentro de la manzana o sección, en caso que se agoten las viviendas, se sustituye la manzana o sección completa, por otra con similares características.

2.1.5. Medición de la Encuesta.

La NENE presenta cambios en el diseño del cuestionario y la manera en que se hacen las preguntas. La nueva encuesta se actualiza a estándares internacionales, teniendo una estructura de secuencia de preguntas que van filtrando la población. Así se determina de manera más precisa la condición de actividad económica de las personas.

Las principales variables de estudio son las características ocupacionales de la población, o sea, si las personas están ocupadas, desocupadas o inactivas. Por medio de estos datos se obtiene principalmente la “**Tasa de Desocupación**”, que como fue mencionado, será nuestro foco de estudio en este proyecto de título.

2.1.6. Definiciones o Conceptos.

La NENE, tiene una clasificación de la población en edad de trabajar (PET) según condición de actividad económica en relación a la fuerza de trabajo. El objetivo es clasificar a todas las personas de 15 años o más, según su situación laboral, basando el enfoque de medición del empleo y desempleo en un conjunto de reglas de nivel internacional.

A continuación se definirán algunos conceptos necesarios, y las grandes categorías en que se clasifica la población, para mejorar el entendimiento del análisis de la condición laboral. Las personas quedarán clasificadas como ocupadas, desocupadas o inactivas, categorías mutuamente excluyente.

1. **Población en Edad de Trabajar (PET):** Población actualmente residente de 15 años y más.

2. **Población Económicamente Activa (PEA) o Fuerza de Trabajo:** Personas en edad de trabajar, que durante la semana de referencia, cumplen los requisitos para ser incluidas en la categoría de ocupados o desocupados.
3. **Ocupados:** Los ocupados son todas las personas en edad de trabajar, que deben cumplir los siguientes criterios:
 - a) Haber trabajado al menos una hora durante la semana de referencia.
 - b) Haber devengado (recibió o recibirá) un pago, ya sea en dinero o en especie, por ese **trabajo**.

El **trabajo** se refiere a cualquier actividad que esté dentro de los límites de producción del Sistema de Cuentas Nacionales (SCN) y cubre toda la producción de mercado, así como algunos tipos de producción de no mercado (productos primarios o procesados para autoconsumo).

Dentro de la categoría de **Ocupados**, se divide en 3 nuevas categorías.

- a) **Ocupados tradicionales:** Son los ocupados que respondieron afirmativamente desde un inicio que trabajaron la semana pasada. Según la encuesta, la sección A del cuestionario, quienes perciban su trabajo como tal, quedarán clasificados en esta categoría.
 - b) **Ocupados No tradicionales:** Son los ocupados que al consultar inicialmente si trabajaron la semana pasada, respondieron que no lo hicieron, y luego en una serie de preguntas más inclusivas y abiertas responden que si trabajaron.
 - c) **Ocupados ausentes:** Son los ocupados que temporalmente no trabajaron durante la semana de referencia por diversas razones, pero que mantuvieron un vínculo estrecho con su empleo y se espera que vuelvan al trabajo después de su periodo de ausencia. Según la encuesta, la pregunta A6 precisa las razones de ausencia de los ocupados en la semana de referencia, que resulta de interés analítico.
4. **Desocupados:** Los desocupados son todas las personas en edad de trabajar, que deben cumplir los siguientes criterios:
 - a) No haber trabajado durante la semana de referencia.
 - b) Haber buscado un empleo en las últimas cuatro semanas, incluyendo la de referencia.
 - c) Estar disponible para trabajar en las próximas dos semanas, posteriores a la de referencia.

Los criterios b) y c) permiten detallar las “presiones de oferta” en el mercado de trabajo. El hecho de buscar trabajo pero no estar disponible implica que, no se está compitiendo por un puesto de trabajo. Por otra parte, estar disponible para trabajar pero sin haber buscado empleo, supone que no participó en el mercado como oferente.

Dentro de la categoría de **Desocupados**, se divide en 2 nuevas categorías.

- a) **Cesantes:** Son todas las personas que habiendo cumplido con los criterios de desocupado, o sea, desea trabajar y ha hecho esfuerzos definidos por conseguir trabajo, tuvo anteriormente un empleo que duró por lo menos 1 mes.
- b) **Buscan trabajo por primera vez:** Son todas las personas que habiendo cumplido con los criterios de desocupado, carecen de experiencia laboral o no han tenido un empleo que durara por lo menos 1 mes.

Definida de esta forma la **Fuerza de Trabajo**, podemos ejemplificar quienes no forman parte de la fuerza de trabajo:

- a) Población menor de 15 años, y
- b) Población económicamente inactiva, es decir: Dueñas de casa; Estudiantes; Jubilados; Pensionados; Personas que reciben ingresos por concepto de rentas u otra inversión, siempre y cuando no trabajen; Personas internadas en instituciones; Inválidas, etc.

5. **Inactivos o Fuera de la Fuerza de Trabajo:** La población inactiva está conformada por todas las personas de la población en edad de trabajar, que no tuvieron un empleo y tampoco lo buscaron y/o estuvieron disponibles para trabajar. Es decir, comprende a todas las personas no clasificadas como ocupadas y desocupadas (aquí surge una innovación de la encuesta, que es la clasificación de los inactivos según razones por las que no participa en el mercado de trabajo). Se distinguen las siguientes razones de inactividad: desaliento; estacionales; temporales; familiares; estudio; jubilación; salud; pensión o montepío; inicio de actividad.

Dentro de la categoría de **Inactivos**, se divide en 3 nuevas categorías.

- a) **Iniciadores:** Son todas las personas de la población en edad de trabajar, que declaran no haber buscado trabajo en las últimas cuatro semanas por la razón de que iniciarán pronto una actividad laboral y que, al mismo tiempo, declara disponibilidad. De estar disponibles, son considerados en la fuerza de trabajo para fines de tasas estandarizadas OCDE (de armonización estadística).
- b) **Inactivos Potenciales Activos:** Son todos aquellos inactivos que, independientemente de la razón a la que está asociada su condición de inactivo, buscaron empleo pero no estarían disponibles para trabajar o no buscaron empleo pero estarían disponibles para trabajar. **Potencial**, se refiere a una presunción de la factibilidad de que una persona que no está presionando el mercado laboral, efectivamente lo haga en algún momento.
- c) **Inactivos Habituales:** Son todos aquellos inactivos que, independientemente de la razón a la que está asociada su condición de inactivo, no buscaron empleo y tampoco estarían disponibles para trabajar.

2.1.7. Principales indicadores de razón.

De acuerdo a las categorías de la condición de actividad de la población en edad de trabajar, se construyen los principales indicadores de razón:

1. **Tasa de Desempleo:** Número de personas desocupadas expresadas como porcentaje de la Fuerza de Trabajo (Ocupados + Desocupados).
2. **Tasa de Ocupación:** Número de personas ocupadas expresadas como porcentaje de la Población en Edad de Trabajar (Ocupados + Desocupados + Inactivos).
3. **Tasa de Participación:** Número de personas en la Fuerza de Trabajo expresadas como porcentaje de la Población en Edad de Trabajar.
4. **Tasa de Desocupación estandarizada:** Número de personas desocupadas más los iniciadores disponibles, expresadas como porcentaje de la Fuerza de Trabajo (Ocupados + Desocupados + Iniciadores Disponibles).

2.1.8. Factor de expansión.

El factor de expansión se interpreta como la cantidad de viviendas en la población que representa una vivienda en la muestra. Lo que se transmite a las personas que viven en la vivienda.

Dicho factor, de acuerdo con el diseño de muestreo estratificado, corresponde al aplicado a una muestra en dos etapas. Este depende muestralmente del número de viviendas que tienen las unidades de la primera etapa (manzanas o secciones) y de la cantidad de estas en cada estrato, y, poblacionalmente del número de viviendas en el estrato y en la sección o manzana seleccionada.

El factor de expansión Teórico $F1_{hi}$, es calculado como el inverso de la probabilidad de selección de cada vivienda. Este factor es constante por sección o manzana y se aplica sobre todas las viviendas y por lo tanto a todas las personas que pertenecen a la selección o manzana **i** del estrato **h**.

El cálculo corresponde a :

$$F1_{hi} = \frac{M_h}{M_{hi} \cdot n_h} \cdot \frac{M'_{hi}}{m_{hi}}$$

Donde:

M_h : Número de viviendas en el estrato h al Censo 2002 (en el caso de secciones) o al año de actualizaciones del marco en el caso de manzanas.

M'_{hi} : Número actualizado de viviendas en la unidad de primera etapa i en el estrato h.

M_{hi} : Número de viviendas en la unidad de primera etapa i en el estrato h al Censo 2002 o al año de actualización del Marco Urbano.

n_h : Número de unidades de primera etapa en la muestra en el estrato h .

m_{hi} : Número de viviendas en la muestra en la unidad de primera etapa i en el estrato h .

El factor de expansión incluye un ajuste de población de personas por estrato, donde cada estrato tiene sus propias estimaciones de proyecciones de población, por sexo y tramo de edad (menores de 15 años y de 15 años y más) con el objeto de aumentar la precisión de las estimaciones y ser consistente con los stocks poblacionales de las proyecciones de población. De esta manera, después de este ajuste, por cada sección o manzana del estrato h , se tienen 4 factores de expansión que se aplican respectivamente para todas las personas dentro de la sección que sean hombres menores (menores de 15 años), hombres mayores (mayores de 14 años o de 15 años o más), mujeres menores y mujeres mayores.

El cálculo del factor Ajustado corresponde a:

$$F2_{hi} = F1_{hi} \cdot \frac{P_h}{\hat{P}_h}$$

Donde:

P_h : Proyección de población total por sexo y tramo de edad en el estrato h considerando como un parámetro para el ajuste.

\hat{P}_h : Número total de personas estimadas con el factor teórico, por sexo y tramo de edad en el estrato h .

$$\text{Con } \hat{P}_h = \sum_i \sum_j F1_{hi} \cdot P_{hij}$$

P_{hij} : Número total de personas (hombres menores, u hombres mayores, o mujeres menores, o mujeres mayores) en la viviendas j , de la unidad de primera etapa (sección o manzana) i en el estrato h .

2.2. Análisis de Correspondencias.

2.2.1. Variables categóricas.

Las variables cualitativas o categóricas son aquellas cuyos valores corresponden a un conjunto de conceptos, atributos o cualidades no numéricas, llamadas categorías, niveles o modalidades. Una propiedad deseable de las categorías es que sean mutuamente excluyentes, o sea, cada individuo se clasifica en una y solo una categoría. La clasificación de las variables cualitativas puede ser según el número de categorías: dicotómicas (solo dos modalidades) o politómicas (más de dos modalidades); o puede ser según la escala de medición de las categorías: nominal (las categorías no poseen un orden natural o jerárquico), ordinal (las categorías poseen un orden natural o jerárquico) o por intervalo (proceden de variables cuantitativas agrupadas en intervalos). Para analizar las variables cualitativas, se deben usar métodos estadísticos apropiados para cada tipo de escala [Aguilera, Ana María. 2005].

En el caso particular de 2 variables categóricas X e Y , con valores $\{x_1, \dots, x_r\}$ e $\{y_1, \dots, y_c\}$ respectivamente (sus correspondientes categorías o modalidades), provenientes de una muestra o población de tamaño n . La forma de representar u organizar los datos es mediante una tabla de contingencia bidimensional, también llamada tabla de doble entrada o tabla cruzada. Las filas están formadas por las categorías de una variable y las columnas por las categorías de otra variable. Comúnmente, las filas están formadas por las categorías de una covariable y las columnas por las categorías de la variable respuesta. El fin es establecer si existe asociación entre las variables categóricas. La tabla de contingencia bidimensional se define generalmente como sigue:

$X Y$	y_1	\dots	y_j	\dots	y_c	Total
x_1	n_{11}	\dots	n_{1j}	\dots	n_{1c}	n_{1+}
\vdots	\vdots		\vdots		\vdots	\vdots
x_i	n_{i1}	\dots	n_{ij}	\dots	n_{ic}	n_{i+}
\vdots	\vdots		\vdots		\vdots	\vdots
x_r	n_{r1}	\dots	n_{rj}	\dots	n_{rc}	n_{r+}
Total	n_{+1}	\dots	n_{+j}	\dots	n_{+c}	n_{++}

Tabla 2.1: Tabla de Contingencia bidimensional.

Donde:

n_{ij} : es el número de individuos o elementos de la muestra que se clasifican simultáneamente en la categoría “ i ” de la variable X y en la categoría “ j ” de la variable Y .

n_{i+} : es el número de individuos o elementos de la muestra que se clasifican en la categoría “ i ” de la variable X , recibe el nombre de frecuencia marginal de la categoría “ i ”.

n_{+j} : es el número de individuos o elementos de la muestra que se clasifican en la categoría “ j ” de la variable Y , recibe el nombre de frecuencia marginal de la categoría “ j ”.

n_{++} : es el número de individuos o elementos total de la muestra, o sea, “ n ”.

Además, la tabla de contingencia bidimensional de frecuencias observadas de la muestra o población, se puede presentar como un arreglo de idénticas dimensiones y modalidades en el cual las celdas o casillas contienen proporciones (probabilidades, también llamada frecuencias relativas) en lugar del número de individuos o elementos n_{ij} . La proporción se obtiene de la forma $\frac{n_{ij}}{n_{++}}$. A esta nueva tabla se le llama distribución de probabilidad de las variables X e Y . Donde se tiene que:

p_{ij} : es la proporción de que un individuo o elemento de la muestra pertenezca a la categoría “ i ” de la variable X y a la categoría “ j ” de la variable Y .

p_{i+} : es la proporción de que un individuo o elemento de la muestra pertenezca a la categoría “ i ” de la variable X . También se llama probabilidad marginal de la categoría “ i ”.

p_{+j} : es la proporción de que un individuo o elemento de la muestra pertenezca a la categoría “ j ” de la variable Y . También se llama probabilidad marginal de la categoría “ j ”.

p_{++} : es la suma de todas las proporciones. Se verifica como cualquier distribución de probabilidad discreta que es igual a 1.

Como ya se mencionó, el fin es establecer si existe asociación entre las variables categóricas y esto se analiza mediante contrastes de hipótesis, habitualmente se utiliza la prueba Chi cuadrado (χ^2) de Pearson, que tiene por hipótesis:

$H_0 : p_{ij} = p_{i+} \cdot p_{+j}$ (No hay asociación entre las variables o independencia).

$H_1 : p_{ij} \neq p_{i+} \cdot p_{+j}$ (Si hay asociación entre las variables o dependencia).

El estadístico utilizado para contrastar las hipótesis es el Chi cuadrado (χ^2) que se define como:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (2.1)$$

que, en la hipótesis de independencia, asintóticamente sigue una distribución χ^2 con $(r - 1) \times (c - 1)$ grados de libertad. Según la notación, e_{ij} es el valor esperado para la fila i y la columna j . Se obtiene de la forma $\frac{n_{i+} \cdot n_{+j}}{n_{++}}$.

El valor p asociado a la prueba Chi Cuadrado corresponde a:

$$P[\chi_{(r-1)(c-1)}^2 \geq \chi_{obs}^2]$$

donde χ_{obs}^2 es el valor observado del estadístico χ^2 , en la muestra.

Por lo tanto, X e Y son independientes si y solo si las probabilidades conjuntas (p_{ij}) son igual al producto de las probabilidades marginales $(p_{i+} \cdot p_{+j})$. Intuitivamente la independencia significa que los valores de una variable no vienen influidos por la modalidad que adopte la otra variable [Aguilera, Ana María. 2005].

En caso opuesto, X e Y son dependientes o existirá relación cuando algunas probabilidades conjuntas sean diferentes del producto de sus probabilidades marginales. Si p_{ij} es superior a este producto, las modalidades i y j se asocian más que en la hipótesis de independencia, o sea, se dice que i y j se atraen. Por el contrario, si p_{ij} es inferior al producto de las marginales, las modalidades i y j se asocian menos que en la hipótesis de independencia, o sea, se dice que existe una repulsión entre estas dos modalidades [Escofier, Brigitte. 1992].

Por otra parte, el estadístico χ^2 para contrastar la independencia puede escribirse en términos de frecuencias relativas:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(nf_{ij} - nf_{i+}f_{+j})^2}{nf_{i+}f_{+j}} \quad (2.2)$$

Luego de un trabajo algebraico se puede expresar como:

$$\chi^2 = n \sum_{i=1}^r f_{i+} \sum_{j=1}^c \frac{1}{f_{+j}} \left(\frac{f_{ij}}{f_{i+}} - f_{+j} \right)^2 \quad (2.3)$$

Siendo cada frecuencia relativa f_{ij} idéntica a la proporción de que un individuo o elemento de la muestra pertenezca a la categoría “ i ” de la variable X y a la categoría “ j ” de la variable Y . Las frecuencias f_{i+} y f_{+j} son las frecuencias relativas marginales de la (fila) categoría i y de la (columna) categoría j , respectivamente.

En la expresión (2.3), la distribución condicionada de las frecuencias relativas de cada fila $(\frac{f_{ij}}{f_{i+}})$, se compara con la frecuencia relativa marginal de las columnas, también llamada distribución media de las filas (f_{+j}) , y cada celda se pondera inversamente a la frecuencia relativa que existe en esa columna. Luego, se suman todas las filas, pero dando a cada fila un peso tanto mayor cuanto mayor es su frecuencia nf_{i+} [Peña, Daniel. 2002].

Bajo esta expresión, si X e Y son independientes, se esperaría que las filas de la tabla de contingencia tendrían una frecuencia relativa condicionada a su marginal similar entre sí y, similar a la frecuencia relativa marginal de las columnas (variable Y).

2.2.2. Análisis de Correspondencia Simple (ACS).

Los orígenes se remontan a 1973, debido a Jean-Paul Benzecri (1932-2019) y colaboradores, de la Universidad de París. El Análisis Factorial de Correspondencias es una técnica descriptiva multivariante que originalmente fue concebida para estudiar las tablas de contingencias, es de gran utilidad puesto que permite representar las categorías de sus filas y columnas de manera sencilla, a través de gráficas. Puede ser vista como el procedimiento que encuentra la mejor representación para ambos conjuntos de datos. El objetivo fundamental es describir e interpretar las relaciones entre las dos variables categóricas involucradas entre sí, dicho en otras palabras, estudiar la desviación de la tabla de contingencia de la hipótesis de independencia. Se construye un diagrama cartesiano basado en la asociación entre las variables, en el que se representan conjuntamente las distintas modalidades de la tabla de contingencia, tal que la proximidad entre los puntos está relacionado con el nivel de asociación entre dichas modalidades. El término correspondencia viene del hecho que la técnica busca las filas (o columnas) que se correspondan en información, o sea, que algunas filas (o columnas) pueden estar proporcionando información equivalente respecto a un conjunto de individuos. En conclusión, el Análisis de Correspondencias es una técnica gráfica que representa información contenida en una tabla de contingencia, que señala la interacción de dos variables categóricas a través de la relación de las filas y de las columnas entre sí. Además, como todo análisis factorial, se utiliza también para obtener un pequeño número de dimensiones (o factores) de los datos, de tal manera que la primera dimensión explique la mayor parte de la asociación total entre filas y columnas, la segunda dimensión explique la mayor parte del resto de la asociación, que no es explicada por la primera dimensión, y así con el resto de las dimensiones, conservando la mayor información posible.

2.2.2.1. Perfiles filas y perfiles columnas.

La tabla de contingencia original no se analiza directamente, dado que la comparación de dos filas (dos columnas) tiene el riesgo de estar influida principalmente por sus respectivas marginales. Por lo tanto, se puede realizar un estudio por medio de las filas o por medio de las columnas, según sea el interés de la variable. Así, para el estudio de las filas, la tabla de contingencia se transforma dividiendo cada término n_{ij} por la correspondiente frecuencia marginal de la categoría “ i ”. La nueva fila se denomina perfil-fila, la nueva tabla será una tabla de probabilidades condicionadas, donde cada probabilidad marginal de la variable fila tendrá el valor de 1.

Por el contrario, para el estudio de las columnas, la tabla de contingencia se transforma dividiendo cada término n_{ij} por la correspondiente frecuencia marginal de la categoría “ j ”. La nueva columna se denomina perfil-columna, la nueva tabla será una tabla de probabilidades condicionadas, donde cada probabilidad marginal de la variable columna tendrá el valor de 1.

$X Y$	y_1	\dots	y_j	\dots	y_c	Total
x_1	$\frac{n_{11}}{n_{1+}}$	\dots	$\frac{n_{1j}}{n_{1+}}$	\dots	$\frac{n_{1c}}{n_{1+}}$	1
\vdots	\vdots		\vdots		\vdots	\vdots
x_i	$\frac{n_{i1}}{n_{i+}}$	\dots	$\frac{n_{ij}}{n_{i+}}$	\dots	$\frac{n_{ic}}{n_{i+}}$	1
\vdots	\vdots		\vdots		\vdots	\vdots
x_r	$\frac{n_{r1}}{n_{r+}}$	\dots	$\frac{n_{rj}}{n_{r+}}$	\dots	$\frac{n_{rc}}{n_{r+}}$	1
Marginal	$\frac{n_{+1}}{n_{++}}$	\dots	$\frac{n_{+j}}{n_{++}}$	\dots	$\frac{n_{+c}}{n_{++}}$	1

Tabla 2.2: Tabla de perfiles fila.

El número $\frac{n_{ij}}{n_{i+}}$ representa la probabilidad de que ocurra la categoría “ j ” dado que posee la categoría “ i ”. Concluyendo que, el perfil fila es la probabilidad condicionada en la categoría “ i ” (de la variable X) sobre el conjunto de modalidades de la variable Y . Para efectos del Análisis de Correspondencia, la frecuencia marginal de la variable Y (ubicada en la última fila de la tabla) pasa a llamarse **perfil fila medio**.

Además, se representa la tabla de perfiles columna. Debido a la simetría por las filas y columnas, el razonamiento es análogo. El número $\frac{n_{ij}}{n_{+j}}$ representa la probabilidad de que ocurra la categoría “ i ” (de la variable X) sabiendo que posee la categoría “ j ” (de la variable Y). En este caso, la frecuencia marginal de X es llamada **perfil columna medio**.

$X Y$	y_1	\dots	y_j	\dots	y_c	Marginal
x_1	$\frac{n_{11}}{n_{+1}}$	\dots	$\frac{n_{1j}}{n_{+j}}$	\dots	$\frac{n_{1c}}{n_{+c}}$	$\frac{n_{1+}}{n_{++}}$
\vdots	\vdots		\vdots		\vdots	\vdots
x_i	$\frac{n_{i1}}{n_{+1}}$	\dots	$\frac{n_{ij}}{n_{+j}}$	\dots	$\frac{n_{ic}}{n_{+c}}$	$\frac{n_{i+}}{n_{++}}$
\vdots	\vdots		\vdots		\vdots	\vdots
x_r	$\frac{n_{r1}}{n_{+1}}$	\dots	$\frac{n_{rj}}{n_{+j}}$	\dots	$\frac{n_{rc}}{n_{+c}}$	$\frac{n_{r+}}{n_{++}}$
Total	1	\dots	1	\dots	1	1

Tabla 2.3: Tabla de perfiles columna.

Se puede estudiar ambas transformaciones en perfiles, pero se prioriza una variable de interés para considerar la respectiva tabla transformada. Normalmente, la tabla de contingencia se presenta en forma de porcentajes, según el perfil de interés.

2.2.2.2. Distancias Chi-Cuadrado (χ^2) e Inercia.

Para estudiar la relación entre dos variables o la relación entre perfiles filas y perfiles columnas, se debe conocer la semejanza entre estos perfiles, es por esto que se define una distancia entre perfiles filas y entre perfiles columnas. Esta distancia se conoce como distancia χ^2 , se define de manera análoga para las filas y las columnas:

$$\text{Distancia entre perfil fila } i \text{ y } l \quad d_{\chi^2} = \sum_j \frac{1}{f_{+j}} \left(\frac{f_{ij}}{f_{i+}} - \frac{f_{lj}}{f_{l+}} \right)^2$$

$$\text{Distancia entre perfil columna } j \text{ y } k \quad d_{\chi^2} = \sum_i \frac{1}{f_{i+}} \left(\frac{f_{ij}}{f_{+j}} - \frac{f_{ik}}{f_{+k}} \right)^2$$

Donde cada valor f_{ij} son las frecuencias relativas de la tabla de contingencia original, análogo a la tabla de distribución de probabilidad de las variables X e Y .

La distancia entre dos filas (dos columnas) depende esencialmente de las diferencias término a término entre los dos perfiles. **Los cuadrados de las diferencias en cuestión, son ponderados por el perfil medio inverso correspondiente**, para equilibrar la influencia de las columnas (filas) sobre la distancia entre las filas (columnas), dicho de otra manera, aumenta los términos con menor frecuencia.

La distancia χ^2 tiene una propiedad fundamental que se conoce como “el principio de equivalencia distribucional”. Según esta propiedad, si dos filas tienen “la misma estructura relativa” $\left(\frac{f_{ij}}{f_{i+}} \right)$ o son proporcionales, y se acumulan en una sola fila, la distancia entre los restantes perfiles filas permanecen invariables, no se alteran. Por simetría, esta propiedad es igualmente aplicable a las columnas. Esta propiedad matemática resulta ser muy práctica, ya que reemplazar dos filas o dos columnas casi proporcionales por la suma de éstas, asegura una cierta invarianza del procedimiento ante agregaciones o desagregaciones irrelevantes u homogéneas de las categorías y, por ende, no modifica sensiblemente los resultados de un Análisis de Correspondencias. Se puede nombrar esta propiedad cuando existen varios conjuntos de modalidades.

Por otra parte, la **Inercia**, corresponde exactamente con la noción estadística de varianza (medida de dispersión de los valores alrededor de la media), es el producto de la ponderación de un punto por el cuadrado de la distancia entre el punto y cualquier otro punto del espacio, que normalmente es el baricentro o perfil medio. La inercia total de una tabla, es la suma de las inercias de los elementos, puntos o perfiles que la componen, sobre el baricentro o perfil medio.

La inercia total de i perfiles fila, asociada a cada j -ésima perfil columna, con respecto al baricentro o perfil fila medio se escribe:

$$\text{Inercia Total} = \sum_i \text{Inercia}(i) = \sum_i f_{i+} \sum_j \frac{1}{f_{+j}} \left(\frac{f_{ij}}{f_{i+}} - f_{+j} \right)^2$$

Al observar la fórmula, podemos asociar la inercia a la distancia chi-cuadrado, diciendo que la inercia total es la suma de las masas de las filas (frecuencia relativa marginal) multiplicada por la distancia χ^2 de los perfiles fila con el perfil fila medio.

Equivalentemente se puede expresar la inercia total en base a los j perfiles columna, asociado a cada i -ésimo perfil fila, con respecto al baricentro (en este caso el perfil columna medio), resultando ambas inercias iguales.

Recordemos que, al momento de estudiar la relación de una tabla de contingencia se realiza frecuentemente por medio del estadístico χ^2 , éste mide la desviación entre el número de individuos o elementos observados y el número de individuos o elementos teóricos (o esperados) en promedio, suponiendo la independencia de las variables.

También podemos asociar lo siguiente, el estadístico χ^2 es idéntico a la inercia total de una tabla de contingencia, excepto por el valor de n . Así, podemos escribir:

$$\chi^2 = n \cdot \text{Inercia total}$$

Por lo que, la inercia total (puede ser por análisis de las filas o análisis de las columnas, dada la simetría) se puede ver como una descomposición de los componentes del estadístico χ^2 en sus fuentes de variación [Peña, Daniel. 2002]. Además, se deduce que la inercia total representa la relación entre las dos variables bajo dos formas diferentes.

2.2.2.3. Las dos nubes de perfiles y sus ajustes.

Para realizar el estudio por medio de los perfiles filas se debe obtener una nube N_I de perfiles fila en R^J y luego se realiza el ajuste de la nube de perfiles fila. Por lo tanto, a cada perfil fila se le asocia un conjunto de J valores numéricos, por lo que se puede representar a cada perfil fila como un punto del espacio vectorial J - dimensional R^J , donde cada dimensión está asociada a una modalidad de la variable columna.

El conjunto de perfiles fila constituye la nube N_I en R^J , cuyo centro de gravedad (o baricentro) G_I tiene por coordenada j -ésima la frecuencia marginal f_{+j} , sobre cada eje j , o sea, es el perfil fila medio, que sirve de referencia en el estudio de las filas de la tabla. Por lo tanto, estudiar en qué medida y de qué manera una categoría “ i ” (perfil fila i -ésimo) difiere del conjunto de la población, implica estudiar la desviación entre el perfil i -ésimo y el perfil fila medio. Entonces, estudiar la dispersión de la nube alrededor de su baricentro, no es más que estudiar la desviación entre los perfiles de las filas y la marginal, o sea, la relación entre las dos variables [Escofier, Brigitte. 1992].

En el espacio R^J , la noción de semejanza entre “modalidades” coincide con la distancia χ^2 . El conjunto de distancias χ^2 constituye la forma de la nube N_I .

El objetivo es aportar imágenes planas, llamadas planos factoriales, que sean un sub espacio de dimensión menor que el espacio R^J , las cuales deben tener como origen el baricentro de la nube N_I y se deben aproximar lo más posible a la nube de puntos N_I situada en espacio R^J , o sea, las imágenes planas deben ser tales que las distancias entre los puntos de la imagen **proyectada** se asemejen lo más posible a las distancias χ^2 entre los puntos de la nube N_I . “Así, el subespacio que mejor se ajusta tendrá la menor suma de las distancias de los perfiles filas de la nube N_I al subespacio encontrado (como los perfiles tienen su propio peso o masa, se entiende que se trata de una distancia χ^2)”, por lo que, conservará el máximo de la información de la nube N_I . La minimización de la suma de las distancias de los perfiles, se obtiene por medio de una descomposición en valores singulares (DVS), un resultado muy útil en teoría matricial. Así, se factoriza dicha matriz de datos de las distancias de los perfiles (centrada y estandarizada) descomponiéndola en sus componentes de los más importantes a los menos importantes, resultando una matriz diagonal (de igual orden que la matriz de datos) con los valores singulares y 2 matrices cuadradas, una en izquierda y la otra en derecha, siendo los vectores propios de la matriz de datos por su transpuesta y los vectores propios de la matriz de datos transpuesta por la matriz de datos, respectivamente. Prácticamente, se busca una sucesión de direcciones privilegiadas de R^J llamadas ejes factoriales, las que tomadas dos a dos, determinan planos factoriales sobre los que se proyecte la nube N_I y cada nueva dirección debe ser ortogonal a las ya halladas. Las rectas o dimensiones obtenidas que mejor se ajustan se denominan ejes principales.

Cada dirección o eje factorial hace máxima la inercia (variabilidad total) proyectada de la nube N_I , entonces, cada eje factorial tiene asociada una inercia y, ésta explica una proporción de la inercia total de la nube. Por lo que, tomando dos dimensiones de la proyección, se habla de la proporción de inercia acumulada explicada por estas dos dimensiones. Este cálculo determina **la Calidad de la Representación de la Proyección**. Dado que la inercia representa la relación entre las dos variables, la inercia de cada eje factorial o dimensión de la proyección de la nube, representa una parte de la relación entre las variables. La inercia asociada a cada eje principal, se llama inercia principal, la que también recibe el nombre de valor propio, ya que se puede calcular como un valor propio de una matriz cuadrada simétrica.

Además, de manera análoga a la inercia total, el estadístico χ^2 se puede definir como: n veces la suma de los valores propios del Análisis de Correspondencias [Cuadras, Carles. 2014]. El número máximo de ejes factoriales o dimensiones es igual al menor número de modalidades de cualquier variable menos uno, pero lo común es que dos o tres dimensiones son suficientes para representar con rigor la asociación entre las dos variables. Por otra parte, la máxima inercia posible coincide con el número máximo de dimensiones. Se sobreentiende que la suma de las inercias principales, asociadas a cada dimensión obtenida, corresponde a la inercia total.

Para realizar el estudio por medio de los perfiles columna, hay que tener en cuenta el papel

simétrico de las filas y las columnas en el Análisis de Correspondencias, todo lo realizado en el análisis de las filas se puede aplicar al análisis de las columnas, teniendo: una interpretación simétrica, la misma inercia total, igual dimensionalidad, el mismo subespacio de mejor ajuste y las mismas posiciones relativas de los perfiles, ya que la forma de crear la nube de perfiles columna N_J en el espacio R^I es de manera idéntica a la nube de perfiles filas. Por lo tanto, poseen las mismas propiedades que el ajuste de la nube en R^J . [14] [21]

“Las dos nubes N_I y N_J son dos representaciones de la misma tabla de contingencia, una por medio de sus perfiles filas y otra por medio de sus perfiles columnas, luego, la relación entre estos dos análisis es reagrupada bajo el término de dualidad” (existencia de dos fenómenos distintos en una misma cosa).

2.2.2.4. Relaciones de transición.

Dada la existencia de esta “dualidad” entre los ajustes de las nubes, además de la idéntica inercia de ellas, cada una tiene sus correspondientes proyecciones de los puntos, por lo que, cada nube tiene un conjunto de coordenadas (factores) de cada modalidad, que pueden ser representadas en función de las coordenadas de las modalidades de la otra nube junto con un factor de escala (λ_i), entonces, estos conjuntos de coordenadas pueden ser representados simultáneamente en un mismo plano factorial.

Así, se definen las coordenadas para los perfiles filas y para los perfiles columna, bajo la definición de relaciones de transición:

$$F_s(i) = \frac{1}{\sqrt{\lambda_s}} \sum_j \frac{f_{ij}}{f_{i+}} G_s(j)$$

$$G_s(j) = \frac{1}{\sqrt{\lambda_s}} \sum_i \frac{f_{ij}}{f_{+j}} F_s(i)$$

Donde:

$F_s(i)$: es la coordenada de la proyección del i -ésimo punto fila, respecto del eje factorial de rango s de N_I ajustado en R^J .

$G_s(j)$: es la coordenada de la proyección del j -ésimo punto columna, respecto del eje factorial de rango s de N_J ajustado en R^I .

λ_s : es la inercia principal, asociada a cada uno de los ejes factoriales de rango s . También llamado autovalor.

Estas ecuaciones, permiten la representación simultánea que se obtiene superponiendo las proyecciones de cada una de las nubes, sobre los mismos planos factoriales (con igual rango de ejes). Así, se denomina una propiedad baricéntrica, la cual especifica que la proyección

de la fila i sobre el eje s es el baricentro de las proyecciones de las columnas j , estando cada columna j ponderada por f_{ij} . Como la simetría es válida en el trabajo de las filas y las columnas, se habla de una doble propiedad baricéntrica. Específicamente, también podemos interpretar esta relación en términos de frecuencia, un punto fila (columna) aparece próximo a aquellas columnas (filas) en las cuales su frecuencia condicional (o perfil) presenta valores más altos y aparece alejado de aquellas columnas en las que su frecuencia condicional tiene valores más bajos.

2.2.2.5. Contribuciones y representación.

Con respecto a los ejes factoriales para las modalidades filas o modalidades columnas se deben definir nuevos términos para analizar la representación de las modalidades en la proyección y realizar comentarios interpretativos.

1. **Las contribuciones absolutas en relación a la inercia principal del eje**, miden la importancia de cada una de las modalidades de las variables analizadas en la construcción de cada eje principal.

Entonces, la contribución absoluta de la fila “ i ” sobre un factor s , expresa la proporción de la inercia de la fila “ i ” que contribuye a la inercia generada por el factor s . La suma de todas las inercias de las proyecciones (o de las contribuciones absolutas) de las filas sobre un factor, es igual a la inercia principal o valor propio generado por el factor. También se define como **la Contribución de una modalidad a la inercia de un eje**. Por lo tanto, se puede definir el porcentaje de variabilidad en el factor s debido a una modalidad de una variable como:

La contribución relativa de la i -ésima fila a la inercia principal del eje s :

$$CR_{is} = \frac{n_{i+}F_s^2(i)}{\sum_{i=1} n_{i+}F_s^2(i)}$$

La contribución relativa de la j -ésima columna a la inercia principal del eje s :

$$CR_{js} = \frac{n_{+j}G_s^2(j)}{\sum_{i=1} n_{i+}F_s^2(i)}$$

2. **Las contribuciones absolutas en relación a la inercia de la fila o de la columna**, miden la importancia de cada eje principal (o factor) para explicar la posición en el espacio ajustado de cada una de las modalidades de las variables.

Entonces, la contribución absoluta del eje s sobre una fila “ i ”, expresa la proporción de la inercia del eje s que contribuye a la inercia del perfil fila “ i -ésimo” (o sea, representa la parte de la distancia al origen de coordenadas explicada por dicho eje). La suma de todas las inercias de los ejes principales o factores sobre la i -ésima fila (o j -ésima columna), es igual a la inercia del perfil fila i -ésimo (o perfil columna j -ésimo). Luego, si dividimos cada componente de la inercia de una fila “ i ” sobre un factor, en la inercia total de la fila “ i ”, obtenemos la contribución relativa (o porcentaje explicado) del eje s sobre la i -ésima fila. Este resultado coincide con el coseno cuadrado del ángulo ω_i formado entre el perfil fila y la proyección de la fila “ i ” sobre el eje s (coseno cuadrado de ω_j en el caso de las columnas). También se define como **la Calidad de representación de una modalidad en un eje o factor**. Por lo tanto:

La contribución relativa (o representación) del eje s a la i -ésima fila, es:

$$CR_s(i) = \cos_s^2(\omega_i)$$

La contribución relativa (o representación) del eje s a la j -ésima columna, es:

$$CR_s(j) = \cos_s^2(\omega_j)$$

También se puede interpretar los cosenos cuadrados como coeficiente de correlación al cuadrado de la fila i -ésima con cada eje principal. Valores de CR_S próximos a 1, dan cuenta de modalidades que influyen o están asociados altamente con respecto al eje “ s ”, por ende, la modalidad se halla muy próxima a dicho eje. Por otra parte, valores bajos de CR_S indican una baja contribución del eje “ s ” en la posición de la respectiva modalidad, por ende, la modalidad será casi perpendicular sobre el eje “ s ”. Por último, la distancia entre 2 modalidades en un plano no es realmente cierta, a menos que estén bien representados en el espacio ajustado.

2.2.2.6. Interpretaciones.

Finalmente, el plano factorial o diagrama cartesiano muestra un punto por cada fila y un punto por cada columna de la tabla de contingencia. Estos puntos son las proyecciones de los perfiles filas y columnas de las nubes N_I y N_J .

Entonces, con respecto a los puntos o modalidades se puede comentar que:

- a) Los puntos más alejados del origen de coordenadas, son los perfiles más diferentes del perfil medio o perfil marginal correspondiente a la otra variable. Por lo tanto, serán modalidades alejadas de la media de las filas o columnas. Por otra parte, las modalidades más cercanas al origen son las más similares al perfil medio de filas o columnas.
- b) Si la posición relativa de dos puntos filas es muy cercana, los perfiles de las dos filas, a través de las columnas, son similares. Análogamente, si la posición relativa de dos

puntos columnas es muy cercana representan columnas con perfiles similares a través de las filas.

- c) Se encuentran en un mismo lado, sobre un factor, una fila i y las columnas j con las que más se asocia, en el lado opuesto aquellas columnas j con las que menos se asocia.
- d) Se encuentran juntos y concentrado en un punto, todos los perfiles filas (o columnas), cuando son similares al perfil medio. Luego, la tabla verifica una relación de independencia y la inercia es nula.

2.2.3. Análisis de Correspondencias Múltiples (ACM).

El Análisis de Correspondencias Múltiples es un método de estudio descrito por “ J ” variables cualitativas, que se encuentran definidas sobre un conjunto de “ I ” individuos. En particular, se entiende como una extensión o generalización del Análisis de Correspondencias simple, dado que se puede estudiar la relación entre varias variables cualitativas (nominales u ordinales) y, desde un punto de vista geométrico y algebraico ambos procedimientos son similares, o sea, utiliza los mismos principios generales del ACS. En la tabla de datos, cada fila es un individuo, observación o caso (un elemento de cualquier naturaleza) de frecuencia uno, y cada columna es una de las variables en estudio. En este método se puede estudiar la relación entre las categorías de las variables según su posición en el mapa perceptual y, se puede conocer la posición de los individuos (objetos o elementos) respecto a las categorías de las variables, pero la representación de los I individuos para este proyecto resulta innecesaria. Una de las aplicaciones más comunes es el tratamiento del conjunto de respuestas, a las preguntas formuladas en una encuesta. Cada pregunta es una variable cuyas modalidades son las respuestas propuestas con elección única.

Al igual que en Análisis de Correspondencias Simples, el objetivo del Análisis de Correspondencias Múltiples es generar mapas perceptuales para estudiar relaciones según distancias entre puntos de dos o más variables categóricas. En ACM los puntos pueden representar a individuos (registros u objetos) y a modalidades de las variables. Por lo que, para el caso de representación de individuos, estos serán similares si asumen aproximadamente las mismas variables, de este modo, la asociación entre variables se presenta ya que serían los mismos individuos que asumen las mismas modalidades de diferentes variables. Los principios generales de ACM son iguales a los de ACS: transformar la tabla de datos en perfiles fila y perfiles columna, por ende asociar cada modalidad un punto en el espacio; ajustar los datos o puntos ponderados por sus perfiles, y; la distancia usada entre perfiles es la χ^2 . Además, el sistema de referencia de los ejes se establece mediante el mismo procedimiento.

2.2.3.1. Tipos de tablas.

Inicialmente, los datos se encuentran en una tabla del tipo “Individuos por Variables”, con: I número de individuos (filas), J número de variables cualitativas (columnas) y, x_{ij} es el

valor o modalidad de la variable j que presenta el individuo i (también se llama codificación condensada). Sin embargo, esta tabla de datos no se trata vía análisis de correspondencias múltiples, ya que las marginales filas o columnas no tienen ningún sentido. Así que, la tabla de datos se debe recodificar para poder realizar ACM. Sin embargo, varios software reconocen la tabla del tipo codificación condensada y realizan una recodificación interna para obtener los resultados.

Una posible codificación es la llamada “**Tabla Disyuntiva Completa (TDC)**” o **matriz binaria**, que se construye modificando la tabla inicial, donde las filas representan de igual forma a los individuos, mientras que, las columnas representan las modalidades de las variables. Entonces, el valor de x_{ik} valdrá 1 si el individuo i posee el atributo de la modalidad k y 0 si no la posee. Por lo tanto, K es la cantidad y conjunto de modalidades de todas las variables en conjunto, además, K_j es el número y conjunto de modalidades de la variable j . Se obtiene $K = K_1 + K_2 + \dots + K_J$. Luego, la TDC será una tabla de tantas variables binarias como modalidades tengan las variables categóricas, como se muestra en la Tabla 2.4.

	Variable 1	...	Variable j	...	Variable J	
Individuos	1 2 3 4 5		k		K	Marginal
1	0 0 0 1 0	...	x_{1k}	...	0 0 0 1	J
\vdots	\vdots		\vdots		\vdots	\vdots
i	1 0 0 0 0	...	x_{ik}	...	0 0 1 0	J
\vdots	\vdots		\vdots		\vdots	\vdots
I	0 0 1 0 0	...	x_{Ik}	...	1 0 0 0	J
Marginal	I_1	...	I_k	...	I_K	

Tabla 2.4: Tabla disyuntiva completa (TDC).

La marginal de individuos, tendrá el valor J (número de variables) para todas las filas por igual, ya que los individuos tendrán valor 1 solo en una modalidad por variable. Por otra parte, la Marginal de variables I_k corresponde a la frecuencia de cada modalidad k para cada variable J , o sea, el número de individuos que respondió aquella modalidad. Se sobre entiende que las modalidades de una misma variable son mutuamente excluyentes, que cada individuo toma el valor de 1 en la modalidad elegida de la variable y 0 en las demás modalidades de la misma variable. La TDC también es conocida como Tabla “ K ”, como Tabla “ Z ” o Tabla “ X ” (según el texto en estudio), definida como una yuxtaposición de $K_1, K_2, K_3, \dots, K_J$; como se aprecia en la Tabla 2.4.

Finalmente, se anticipa que el Análisis de Correspondencias Múltiples no es más que un análisis de correspondencias sobre la matriz binaria, la cual proporcionará coordenadas tanto para las filas (individuos) como para las modalidades. Se recuerda que el interés del trabajo es representar solo las modalidades.

Otra posible codificación es la llamada “**Tabla de contingencia Burt**” o solo tabla o

matriz Burt, es una yuxtaposición de tablas de contingencia entre las variables analizadas tomadas dos a dos, y cada individuo aparece en ella J^2 veces, o sea está formada por J^2 bloques o subtablas. Esta tabla se puede obtener por la multiplicación matricial de la TDC, de la forma: $B = K'K$, la matriz Burt (B) se obtiene de matriz K (TDC) transpuesta por matriz K . Las subtablas o bloques que contienen la diagonal, cruzan cada variable consigo misma, por lo tanto cada subtabla contiene ceros fuera de la misma diagonal de cada bloque, ya que dos o más modalidades de una misma variable no pueden ser seleccionadas simultáneamente y, en la diagonal contiene la frecuencia de individuos que posee el atributo de la modalidad k , es decir, la diagonal es equivalente a la marginal de variables I_k de la TDC.

Esta tabla es análoga a una matriz de correlaciones, dado que es simétrica por bloques y toma el conjunto de las relaciones o todas las combinaciones 2 a 2 entre las variables analizadas, por lo que contiene todas las tablas de contingencia simple entre las variables, pero doblemente, por sobre y bajo la diagonal principal. La cantidad de cruzamientos distintos fuera de la diagonal son $J(J - 1)/2$, pero transpuestos en ambos lados de la diagonal.

		Variable 1	...	Variable j	...	
		1	...	k	...	K
Variable j	1	$\begin{matrix} \cdot\cdot \\ J_1 \times J_1 \\ \cdot\cdot\cdot \end{matrix}$...	$\begin{matrix} \cdot\cdot & \vdots \\ J_1 \times J_j & \\ \cdot\cdot & \vdots \end{matrix}$...	
	\vdots		\vdots		\vdots	
	k	$\begin{matrix} \cdot\cdot & \vdots \\ \cdot\cdot & I_k \cdot\cdot \\ \cdot\cdot & \vdots \cdot\cdot \end{matrix}$
\vdots		\vdots		\vdots		
Variable l	q	$\begin{matrix} \cdot\cdot & \vdots \\ \cdot\cdot & I_{qk} \cdot\cdot \\ J_l \times J_1 & \end{matrix}$...	\vdots	...	$\begin{matrix} \cdot\cdot \\ \cdot\cdot & \vdots \\ \cdot\cdot & \vdots \end{matrix}$
K		\vdots	...	$\begin{matrix} \cdot\cdot \\ \cdot\cdot & \vdots \\ \cdot\cdot & \vdots \end{matrix}$
Marginal		$J I_k$				

Tabla 2.5: Tabla de Burt (B).

El valor de I_{qk} corresponde a la frecuencia o número de individuos que poseen la modalidad q de la variable l y la modalidad k de la variable j . Por último, al igual que la matriz binaria, el ACM se basa, en realizar un análisis de correspondencias sobre la llamada Tabla o Matriz de Burt y, por su simetría, las soluciones de filas y columnas son idénticas.

2.2.3.2. Perfiles, Distancias, Inercia, Relaciones de Transición.

En esta sección se muestran diferentes medidas relacionadas al ACM que son homólogas a las medidas presentadas en ACS. Se presentan a continuación, y en primera instancia, diferentes resultados para el ACM de una Tabla Disyuntiva Completa (TDC).

Con respecto a los perfiles de una TDC, de manera análoga que en ACS, estos se definen en relación a su correspondiente marginal. Así, los perfiles filas se obtienen del valor original de la TDC dividido por la correspondiente marginal columna (que es constante). Por el contrario, los perfiles columnas se obtienen tomando el valor original de la TDC y se divide por la marginal fila correspondiente. Luego los perfiles junto a sus baricentros son de la forma:

Perfil del individuo $i = \frac{x_{ik}}{J}$.

El cual simplemente tomará valores 0 o $1/J$, según si la modalidad k es elegida por el individuo. De esta forma cada perfil fila es un punto de la nube $N(I)$ en el espacio R^K .

El perfil fila medio o baricentro de la nube $N(I) = \frac{I_k}{IJ}$, según la modalidad k .

Perfil de la modalidad $k = \frac{x_{ik}}{I_k}$.

En resumen, tomará valores 0 o $1/I_k$, si el individuo i elige la modalidad k . Así, con cada perfil columna se genera la nube de modalidades $N(K)$ en el espacio R^I .

El perfil columna medio o baricentro de la nube $N(K) = \frac{1}{I}$.

Se observa que el baricentro es constante para los perfiles de las modalidades, en particular, un perfil de modalidad k se asemeja cada vez más al perfil medio cuanto mayor es la frecuencia de la modalidad k elegida por los individuos.

Con respecto a las Distancias de los perfiles de una TDC, se encuentran definidas para filas y columnas, sin embargo, se debe tener presente que las filas están compuestas por individuos (registros o casos) y las columnas por cada una de las modalidades de las J variables. Entonces, las distancias entre individuos y entre modalidades están definida por:

$$\text{Distancia entre dos individuos } i \text{ y } l. \quad d^2(i, l) = \frac{1}{J} \sum_k \frac{1}{I_k} (x_{ik} - x_{lk})^2$$

$$\text{Distancia entre dos modalidades } k \text{ y } h. \quad d^2(k, h) = \sum_i I \left(\frac{x_{ik}}{I_k} - \frac{x_{ih}}{I_h} \right)^2$$

Se debe tener en cuenta que, la marginal de los individuos es constante y, es la cantidad de variables J . Por otra parte, la marginal de modalidades es su respectiva frecuencia absoluta o

la cantidad de individuos que seleccionó esa modalidad. Además, los valores x_{ik} representan 0 y 1.

En particular, dos individuos se asemejan si han elegido globalmente las mismas modalidades, así, su distancia disminuye y estos serán cercanos. Luego, si las modalidades entre los individuos difieren, o sea, no han respondido de la misma manera, su distancia aumenta y los individuos estarán alejados. Cada modalidad interviene con la inversa de su frecuencia. La presencia de una modalidad rara o de poca frecuencia, tiende a alejar a aquellos individuos que la seleccionaron con los que no la seleccionaron.

Por otro lado, dos modalidades de una misma variable están lógicamente bastante alejadas una de otra, aunque, si éstas han sido elegidas por grupos semejantes de individuos, o sea, presentan un comportamiento similar en las otras variables, las modalidades estarán próximas.

Además, dos modalidades de distintas variables están alejadas para quienes poseen una y sólo una de las dos modalidades, o sea, su distancia aumenta. Luego, si las modalidades son elegidas de igual manera por los mismos individuos, su distancia disminuye y se encontrarán cercanas o coincidirán. Las modalidades raras o de frecuencia baja estarán siempre alejadas del centro de gravedad o baricentro, y de todas las demás.

Con respecto a la Inercia de una TDC, se encuentra definida para individuos y modalidades, sin embargo en ACM se hace necesario definir la inercia de variables. Por lo que, se define la Inercia de una modalidad k , la inercia de una variable j y la inercia total de la nube, como:

$$\begin{aligned} Inercia(k) &= \frac{I_k}{IJ} \sum_i I \left(\frac{x_{ik}}{I_k} - \frac{1}{I} \right)^2 = \frac{1}{J} \left(1 - \frac{I_k}{I} \right) \\ Inercia(j) &= \sum_{K_j} Inercia(k) = \sum_{K_j} \frac{1}{J} \left(1 - \frac{I_k}{I} \right) = \frac{1}{J} (K_j - 1) \\ Inercia Total &= \sum_k Inercia(k) = \sum_k \frac{1}{J} \left(1 - \frac{I_k}{I} \right) = \frac{K - J}{J} \end{aligned}$$

Con respecto de la inercia de una modalidad podemos decir que cuando menos frecuencia posea, mayor será su inercia. De esta forma, la modalidad aparecerá más alejada del origen. La máxima inercia se obtiene por una modalidad que no es elegida por nadie. Por lo que, es frecuente ver los primeros factores de un ACM representados casi exclusivamente por las modalidades de menor frecuencia o también llamadas modalidades raras. Se intenta en la práctica evitar las modalidades elegidas poco frecuentes, ya que no conviene incluirlas en el análisis.

Sobre la inercia de la variable, ésta aumentará si el número de modalidades aumenta. La inercia mínima de una variable será $1/J$, se obtiene con $K_j = 2$. Referente a la Inercia To-

tal, en ACM no interviene en la interpretación y no tiene ningún significado estadístico. Sin embargo, se utiliza el valor $1/J$ como umbral para observar inercias principales superiores a dicho valor, así dichos ejes asociados serán interesantes para interpretar en el ACM [Greenacre, 2008].

Con respecto a las Relaciones de Transición de una TDC, de forma análoga que en ACS, se encuentran definidas para filas y columnas de la forma:

$$\begin{aligned} F_s(i) &= \frac{1}{\sqrt{\lambda_s}} \sum_k \frac{x_{ik}}{J} G_s(k) \\ G_s(k) &= \frac{1}{\sqrt{\lambda_s}} \sum_i \frac{x_{ik}}{I_k} F_s(i) \end{aligned}$$

Con estos resultados, en específico el correspondiente a las modalidades, $G_s(k)$ es la coordenada de la proyección de la k -ésima modalidad, respecto al eje factorial de rango s de la nube de modalidades. Complementando, se interpreta que la modalidad k está situada en el baricentro de los individuos que la han elegido, excepto por el coeficiente de dilatación $1/\sqrt{\lambda_s}$. Recordemos que λ_s es la inercia principal asociada al eje factorial s , o autovalor. Interpretación análoga referente a la coordenada de los individuos $F_s(i)$.

Con respecto al ajuste del ACM, se descompone la nube de puntos o inercia total del espacio de individuos o de modalidades, tal que busca aquellos factores que se encuentren altamente correlacionados con todos los grupos de modalidades, de esta forma el factor maximizará la inercia explicada. El siguiente factor cumple la condición de ortogonalidad sobre los factores anteriores y, explicando la mayor cantidad de inercia que no fue explicada por el factor anterior. El ajuste óptimo del ACM, como se mencionó, queda en evidencia cuando el factor está correlacionado con todos los grupos de modalidades, esto se produce cuando la dispersión de las distintas modalidades de una misma variable es máxima y, a la vez, ocurre que la dispersión de los individuos (o sea, la distancia al cuadrado) con respecto a sus modalidades elegidas es mínima.

2.2.3.3. Contribuciones, Medidas de Discriminación, Variables en ACM.

Con respecto a las contribuciones de una TDC, son completamente equivalentes y casi las mismas a las de ACS, mas se debe definir una con respecto a las variables.

La contribución absoluta de la variable J en relación a la inercia principal del factor s , mide la importancia de la variable J en la construcción del factor o eje principal y, expresa la proporción de la inercia que contribuye a la inercia principal generada por el factor s , está formada por la suma de las contribuciones de las respectivas modalidades sobre el factor en cuestión, se define como:

$$CA_{K_j s} = \sum_{K_j} \frac{I_k G_s^2(k)}{IJ}$$

La suma de todas las contribuciones absolutas o inercias de las proyecciones de las variables sobre el factor s , es igual a la inercia principal. Por lo tanto, la inercia total del factor s se define como:

$$\lambda_s = \sum_k CA_{K_j s}$$

Por último, se puede definir el porcentaje de variabilidad en el factor s debido a una variable J , entonces se llama contribución relativa de la variable J a la inercia principal del eje s , y se define como:

$$CR_{K_j s} = \frac{CA_{K_j s}}{\lambda_s}$$

Por otra parte, se define la contribución absoluta en relación a la inercia de la variable J , mide la importancia de cada factor para explicar la inercia de la variable. En este caso, expresa la proporción de inercia del factor s que contribuye a la inercia de la variable J . La suma de todas las contribuciones absolutas de cada factor sobre la variable J , es igual a la inercia de la variable. Luego, al dividir la contribución absoluta de un factor en la suma de todas las contribuciones absolutas sobre los factores de la variable J , se obtiene la contribución relativa del factor s sobre la variable J , y se define por:

$$CR_{s K_j} = \frac{CA_{K_j s}}{C_{K_j}}$$

Con C_{K_j} como la contribución absoluta de la variable J sobre todos los factores. También se define, la calidad de reconstitución de la inercia de la variable J por m ejes, como la suma de las contribuciones relativas de aquellos m ejes sobre la variable [Luis Joaristi, 1999].

Con respecto a las Medidas de Discriminación o representación de variables, a pesar de las nuevas definiciones de contribuciones absolutas (inercias) o relativas (porcentuales) de la variable J a la inercia principal del eje s y de la contribución en forma absoluta o relativa de cada factor sobre la variable, resulta mejor obtener una interpretación de las inercias principales junto a su descomposición como correlaciones, en vez de como se presenta en ACS. Así, se debe conceptualizar cada variable al momento de la interpretación, de manera que usaremos la Contribución absoluta $CA_{K_j s}$, la cual se obtiene por medio de sus modalidades que la conforman. Sin embargo, esta contribución está limitada, ya que su inercia comprende valores entre 0 y $1/J$, por lo que, el objeto de análisis de la relación de los factores con las variables en estudio, se hace por medio de la Razón de Correlación (Escofier)

o Valores de Discriminación (Greenacre) o simplemente como aparece en algunos software estadísticos, “**Medidas de Discriminación**”. Entonces, la Razón de Correlación o medida de discriminación se define como:

$$\eta^2(F_s, K_j) = J \cdot CA_{K_j s}$$

La que representa la varianza de las puntuaciones de las modalidades de la variable J en el factor s , la cual debe ser máxima para obtener los resultados óptimos. De esta forma, se toma en cuenta las variables que participan en la definición de cada factor, según las modalidades que aportan a cada eje factorial. Esta medida de discriminación de la variable J está comprendida entre 0 y 1. Cuando es cercana a 1, los individuos que han elegido una misma modalidad están muy reagrupados y las modalidades están separadas unas de otras, en este caso se dice una relación muy fuerte entre la variable y el factor, caso contrario ocurre si el valor es más cercano a 0. Se puede obtener las inercias principales asociadas a cada factor, al realizar una sumatoria de todas las medidas de discriminación de las variables y despejarla, obtenemos las medias de las correlaciones al cuadrado para cada factor, la que coincide con la inercia principal de cada factor.

Es posible visualizar las medidas de discriminación mediante gráficos, donde las variables tendrán por coordenadas sobre el eje s_1 su razón de correlación del factor s_1 y, otra coordenada sobre el eje s_2 su razón de correlación del factor s_2 . Siendo los factores s_1 y s_2 los ejes de abscisas y ordenada, entonces cada variable está representada por un punto. Luego, si una variable J está más cercana al factor s_1 , dado que su valor es más cercano a 1, se dirá que la variable J está muy relacionada con el Factor s_1 . De esta forma, facilitará una selección de variables más relacionadas al factor para apoyar mayormente su interpretación y, además, proporcionará una visualización de las proximidades entre variables, o sea, reflejará sus semejanzas. Resulta interesante comenzar el análisis de los resultados de un ACM mediante la Razón de Correlación, ya que pone de manifiesto las variables más relacionadas con cada uno de los factores [Escofier, 1992].

Con respecto al caso especial de las variables en ACM, en la definición de la matriz binaria (sección 2.3.1) se observa que las variables van introducidas implícitamente, por medio de sus modalidades. A pesar de esto, la nube $N(K)$ de todas las modalidades (en el espacio R^I , análogo que en ACS) se puede descomponer en subnubes de modalidades de una misma variable. A raíz de esto, hemos podido observar definiciones con respecto a las variables. Las subnubes tienen el mismo centro de gravedad o baricentro que la nube global, el baricentro es $1/I$. Por lo tanto, el conjunto de modalidades de una sola variable está centrado sobre todas las subnubes del resto de las variables, y se confunde con el de la nube global de modalidades.

En ACM, una variable de k modalidades es representada en un subespacio de dimensión igual a $k - 1$, entonces, para obtener una representación exacta de las k modalidades de una misma variable es necesario al menos $k - 1$ factores. Siguiendo esta analogía para cada una de las variables, obtenemos que la cantidad máxima de factores de un ACM es la cantidad total de modalidades menos el número de variables ($K - J$). Además, al dividir esta diferencia en J , obtenemos la Inercia Total para una TDC, renombrándose también como la media

del número máximo de factores. Existen algunas consecuencias dada la proyección de las modalidades de una variable:

- a) El porcentaje de inercia asociado a cada factor, en especial al primero es prácticamente débil cuando hay variables con muchas modalidades.
- b) Una variable que posea una medida de discriminación alta sobre un factor, o sea, que esté muy relacionada a un factor, es imposible que represente así de bien todas sus modalidades sobre dicho factor.
- c) Si la cantidad de individuos o registros es muy grande, no resulta útil considerar muchas modalidades de una sola variable.

2.2.3.4. Resultados de ACM para una matriz de Burt, comparación y similitudes con TDC.

Como se anticipó, el ACM sobre una matriz Burt no es más que el ACS de la matriz Burt, se puede decir que el análisis de correspondencias aplicado a una TDC es equivalente al aplicado a la matriz de Burt (B) y produce los mismos factores, esto se debe a la relación directa entre las tablas ($B = K'K$), por lo tanto, conducen al mismo resultado y los dos análisis son similares, excepto un ligero cambio de escala en los ejes, dado que la matriz Burt genera coordenadas en una escala reducida en comparación con la TDC [Greenacre, 2008].

Además, sobre la relación entre las tablas, la k -ésima fila de la matriz Burt es la suma de los individuos de la TDC que eligieron la modalidad k . Geométricamente hablando, esto es que en el espacio correspondiente (R^K) el perfil de la modalidad k de la matriz Burt, se encuentra en el baricentro de los perfiles individuos de la TDC que poseen la modalidad k . Recordemos que la TDC y la matriz Burt tienen la misma marginal sobre las modalidades, excepto por el coeficiente de J . Luego, la estructura, las distancias, los baricentros, entre otras, del espacio R^K en el ACS de TDC y matriz B , es la misma y, se sitúan en el mismo espacio euclídeo. Específicamente, los ejes de inercia de la nube de los perfiles individuos de la TDC, se mezclan y no se distinguen con los de sus baricentros, vistos como filas de la matriz Burt [Escofier, 1992].

Para ver el detalle de los elementos o medidas relacionadas al ACM de una matriz Burt (como marginales, perfiles, masas, entre algunas, y su forma matricial) se aconseja ver [19]. Además, si se busca una explicación geométrica más detallada de la equivalencia entre ambos análisis (como análisis y ajuste de las nubes filas y columnas, dualidad, demostraciones, etc) se aconseja ver [14]. Por último, si busca explicaciones más pragmáticas se aconseja ver [21].

La Inercia Total sobre la matriz de Burt, se define como:

$$IT^B = \sum_s \lambda_s^B = \frac{1}{J^2} \left(\sum_{j \neq j} I(J, J) + (K - J) \right)$$

Donde, $I(J, J)$ es la inercia de cada subtabla o bloque $J \times J$ de la matriz Burt [Cuadras, 2014], y λ_s^B es la inercia principal asociada al eje factorial s de la matriz de Burt, o valor propio de Burt. Se puede observar la inercia total de Burt como la media de las inercias de cada uno de los bloques de la matriz Burt. Es más, puede llegar a ser de interés interpretar la inercia total, dado que en una tabla de contingencia la inercia es proporcional al χ^2 de independencia, por lo tanto, se puede mostrar que la inercia total es igual a la suma de los χ^2 de independencia de cada una de las subtablas o bloques de la tabla Burt.

Las inercias principales asociadas al factor s de la matriz Burt son iguales a los cuadrados de las inercias principales de la TDC, así:

$$\lambda_s^B = (\lambda_s^K)^2$$

Entonces, los valores propios de Burt disminuyen más rápido que los valores propios del análisis de la TDC, como consecuencia, los porcentajes de inercia siempre serán mayores en el ACM de la matriz Burt. Por tanto, hay que ser prudente al retener factores [Joaristi, 1999] y como se mencionó se retienen aquellos factores cuya inercia principal sea mayor que $1/J$. Sin embargo, la inercia de Burt está algo sesgada por las subtablas o bloques de la diagonal, lo que explicaría, al igual que para la inercia de la TDC, bajos porcentajes de inercia de los ejes del análisis de la matriz de Burt [Greenacre, 2008].

Ya sabemos que la matriz de Burt tiene dimensiones más pequeñas que la TDC, por ende las coordenadas factoriales definidas no son iguales, pero sí homotéticas de razón $\sqrt{\lambda_s}$. Por lo tanto, las coordenadas factoriales de la matriz Burt, tienen la siguiente relación con las relaciones de transición o coordenadas factoriales del análisis de una TDC:

$$G_s^B(k) = \sqrt{\lambda_s^K} G_s^K(k)$$

Con respecto a la Inercia Total de la TDC, ya se ha visto que tiene poco sentido, los porcentajes de inercia explicados son bajos, o sea, son inercias pesimistas. Sin embargo, esta situación se puede corregir, y conviene medir la tasa de inercia, utilizando las inercias principales de la matriz Burt o aplicando la siguiente **“Corrección de Benzecri”** para medir la inercia principal: [19] [21]

$$\lambda^c = \left(\frac{J}{J-1} \right)^2 \left(\sqrt{\lambda_B} - \frac{1}{J} \right)^2$$

Para calcular el porcentaje de inercia explicada y obtener las nuevas tasas de inercia con las inercias principales corregidas, se utiliza como inercia total la suma de todas las inercias principales corregidas. Los inercias principales originales y seleccionadas para corrección son las inercias principales retenidas, o sea, las mayores o iguales a $1/J$.

A pesar de esto Greenacre propone corregir de igual forma las inercias de la matriz Burt, observando la relación entre ambas ($\sqrt{\lambda_B} = \lambda_K$), se puede usar la misma corrección. Sobre la Inercia Total corregida, obtiene la media de los bloques de la matriz Burt pero sin incluir los

bloques de la diagonal, dado que al incluir las inercias de los bloques diagonales (la variable con ella misma) conduce a una sobreestimación de la Inercia Total de Burt.

2.2.3.5. Otras características, propiedades, conclusiones e interpretaciones.

Esta última sección está destinada a mostrar algunas características y propiedades que se derivan de la teoría entregada en este capítulo. Se destaca lo siguiente:

- a) El ACM es comparado con la herramienta estadística Análisis de Componentes Principales (abreviado como ACP), dado que puede graficar tanto a individuos como a sus variables (por medio de sus modalidades de forma binaria), se le ha llamado el ACP de las variables categóricas. Sin embargo, poseen algunas diferencias, entre ellas es que los primeros ejes del ACM explican una pequeña parte de la variabilidad total, a diferencia del ACP.
- b) Se dice que el ACS es un caso particular del ACM, ya que el ACS representa dos variables categóricas. Además, se pueden obtener los mismos resultados de un ACS, si se trabajan dos variables como una TDC y como una tabla Burt.
- c) Sobre las relaciones de transición, ambos tipos de tabla de datos tienen las coordenadas para las modalidades, obviamente, y también para los individuos. Pero para efecto de este trabajo se dio principal énfasis en la estructura de las modalidades.
- d) Sobre el número de modalidades, si éstas aumentan manteniendo invariante la cantidad de variables, o sea, se subdividen nuevas categorías o se crean nuevas de las ya existentes, la inercia de la nube $N(K)$ aumentará. En consecuencia, la importancia relativa de cada factor disminuirá en general, sin aportar nada nuevo [Joaristi, 1999].
- e) Sobre el conjunto de modalidades de una variable, dado que está centrado junto con todas las subnubes de las variables y, coinciden con el origen de coordenadas de la proyección de la nube de modalidades, las variables dicotómicas tendrán sus modalidades ubicadas en forma opuesta con respecto al origen.
- f) Sobre la codificación de variables cuantitativas, al agrupar mucho las clases o disminuirlas en exceso, se corre riesgo de perder información, ya que se reagrupan individuos cada vez más diferentes. Tampoco resulta útil superar 8 clases de modalidades. Comúnmente, agrupar una variable cuantitativa en 4 o 5 modalidades es suficiente [Escofier, 1992].

Finalmente, las interpretaciones del ACM están relacionadas y se desprenden de las distancias (sección 2.3.2) entre modalidades. Las principales interpretaciones, sobre algunas medidas y en particular sobre los resultados de un gráfico conjunto, son las siguientes:

- a) Una mención sobre los individuos (que no es de interés para el trabajo), dos individuos serán parecidos o semejantes si tienen casi las mismas modalidades. Entonces, los individuos estarán próximos.
- b) La lejanía de una modalidad del origen es inversamente proporcional a su frecuencia (a diferencia del ACS). Entonces una modalidad con poca frecuencia se ubicará muy alejada del origen, en consecuencia tendrá una mayor inercia y será muy influyente. Por el contrario, una modalidad de mucha frecuencia se ubicará más cercana al origen.
- c) La proximidad entre dos modalidades de una misma variable, por construcción se rechazan entre ellas. Sin embargo, llegan a ser parecidas o semejantes, por ende estar cercanas, ya que los grupos de individuos que las han elegido tienen casi el mismo comportamiento en las otras variables.
- d) La lejanía entre modalidades, representa individuos que han elegido distintos grupos de modalidades, por lo tanto menor asociación existe entre ellas.
- e) La proximidad entre modalidades de variables diferentes, estarán cercanas o próximas las modalidades que han sido elegidas por los mismos individuos o presentan individuos parecidos. Así, las modalidades estarán asociadas.



Capítulo 3

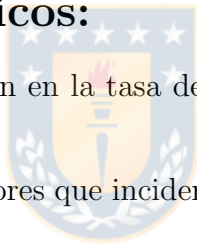
Objetivos.

3.1. Objetivo General:

Modelar la empleabilidad e identificar una estructura en la tasa de desempleo de la Región del Biobío.

3.2. Objetivos Específicos:

1. Identificar los factores que inciden en la tasa de desempleo en la Región del Biobío y sus Provincias.
2. Analizar la estructura de los factores que inciden en la tasa de desempleo en la Región del Biobío y sus Provincias.



Capítulo 4

Materiales y Métodos.

Para realizar este proyecto se recopilaron las bases de datos de la nueva encuesta nacional de empleo (NENE) del trimestre móvil Enero-Febrero-Marzo del año 2010 hasta el último trimestre móvil del año 2015, solicitadas al Instituto Nacional de Estadística. Las bases se encontraban en extensión del software SPSS, se recopilaron un total de 71 bases de datos, una por cada trimestre móvil.

Se procedió a concatenar todas las bases de datos en una sola, comenzando de manera anual y luego los 6 años, para trabajar de manera más eficiente. Se obtuvo una base de datos país desde el año 2010 al 2015 con 7.754.649 registros. A continuación, se eliminaron de la base las personas menores de 15 años, ya que no corresponden a la Población en Edad de Trabajar (PET). Además, las personas inactivas (fuera de la Fuerza del Trabajo) e iniciadoras también se eliminaron, ya que el objetivo es trabajar con los registros necesarios para obtener la Tasa de Desempleo y según definición de Indicadores de razón quedan excluidos; resultando 3.584.034 registros para el país. Por último, se separó una base de datos exclusivamente de la Región del Biobío de los 6 años en cuestión, resultando con 463.008 registros, correspondiente a la Fuerza de trabajo.

Para mejorar el entendimiento sobre la tasa de desempleo y cómo se obtiene en cada trimestre móvil, se replicaron todos los índices de desempleo de la Región del Biobío del trimestre móvil Enero-Febrero-Marzo del año 2010, hasta el trimestre móvil Noviembre-Diciembre-Enero (2016) del año 2015, resultando idénticos a los presentados por el INE [1]. También se replicaron las tasas de desempleo anual para la Región del Biobío, éstas se calculan promediando las tasas de desempleo de los siguientes 4 trimestres móviles: Enero-Febrero-Marzo; Abril-Mayo-Junio; Julio-Agosto-Septiembre; y Octubre-Noviembre-Diciembre; o simplemente de los meses centrales Febrero, Mayo, Agosto y Noviembre.

Las tasas de desempleo fueron obtenidas utilizando el factor de expansión de cada registro, el cual viene como una variable numérica en todas las bases de datos, recordemos que “el factor de expansión corresponde a la cantidad de personas en la población que representa un registro o un individuo en la muestra”.

Para calcular la tasa de desempleo se necesita el número de personas desocupadas y el número

de personas ocupadas. Por ejemplo, para calcular el número de personas desocupadas de un trimestre móvil, se determina como:

$$\sum_i (\text{factor de expansión}_i[\text{desocupado}_i])$$

De manera análoga, se puede calcular el número de personas ocupadas. Luego, el número de personas desocupadas se expresa como porcentaje de la Fuerza de Trabajo (número de personas ocupadas más número de personas desocupadas).

Finalmente, como la información de un mes se repite en las bases de datos de 3 trimestres móviles, se decidió observar los 4 trimestres móviles utilizados para obtener el promedio del año 2015 y seleccionar dos de ellos. El criterio para la selección de los dos trimestres móviles fue: el trimestre de mayor y de menor índice de desempleo. Se obtuvo el trimestre móvil de Julio-Agosto-Septiembre de 2015 con un índice de desempleo del 7,9% (el más alto), y el trimestre móvil de Octubre-Noviembre-Diciembre de 2015 con un índice de desempleo del 6,9% (el más bajo). Se destaca que los índices de desempleo fueron obtenidos sin incluir la Provincia de Ñuble, actualmente Región de Ñuble. Por lo tanto, se realizaron análisis para los 2 trimestres móviles, a modo de comparación y para evitar información mensual repetida.

En la realización de este proyecto se utilizó, principalmente el software SPSS, donde se obtuvieron tablas de frecuencias para las principales variables trabajadas con objeto de describir y caracterizar ambos trimestres móviles, se obtuvieron tablas de contingencia y se buscó relaciones de la variable “Condición de Actividad” (ocupado, otros ocupados y desocupado) con el resto de variables trabajadas a modo de observar influencias en la variable de interés por las variables trabajadas, se obtuvieron Análisis de Correspondencia Simple para identificar factores que inciden en la Condición de Actividad, en específico para describir las relaciones entre variable de interés y el resto de variables, e interpretar las modalidades que inciden en la Condición de desocupado; por último, se obtuvieron Análisis de Correspondencias Múltiples en ambos meses centrales para analizar la estructura de los factores que inciden en el desempleo, de forma equivalente se obtuvieron Análisis de Correspondencias Múltiples para cada una de las Provincias en ambos meses centrales para analizar la estructura de los factores que inciden en la tasa de desempleo por provincia.

Las variables utilizadas fueron 12, todas extraídas de la encuesta y algunas variables re-categorizadas para evitar categorías nulas. La variable dependiente o variable de interés se describe a continuación:

Condición de Actividad:

1. **Ocupado**, que corresponde a: Ocupado tradicional.
2. **Otros Ocupados**, que corresponde a: Ocupado no tradicional y Ocupado ausente.
3. **Desocupado**, que corresponde a: Cesante y Busca trabajo por primera vez.

Las variables independientes se describen a continuación:

Provincia:

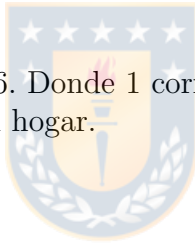
1. Concepción.
2. Arauco.
3. Biobío.

Tipo de Estrato:

1. Grandes centros urbanos o ciudades (abreviada como CD).
2. Restos de áreas urbanas (abreviada como RAU).
3. Área rural (RURAL).

Número de Personas en el Hogar:

Variable discreta, toma valores de 1 a 6. Donde 1 corresponde a una personas en el hogar y 6 corresponde a 6 o más personas en el hogar.



Edad:

Variable numérica categorizada de la siguiente forma:

1. De 15 a 24 años.
2. De 25 a 34 años.
3. De 35 a 44 años.
4. De 45 a 54 años.
5. De 55 a 64 años.
6. De 65 o más años.

Sexo:

1. Hombre.
2. Mujer.

Años de Estudio:

Variable numérica creada de las variables “nivel más alto aprobado” y “curso más alto aprobado”. Luego, categorizada de la siguiente forma:

1. De 0 a 8 años de estudio.
2. De 9 a 12 años de estudio.
3. De 13 a 18 años de estudio.
4. De 19 o más años de estudio.

Estado Conyugal:

1. Casado(a).
2. Conviviente.
3. Soltero(a).
4. Otros. Se agrupó viudo(a), separado(a)/anulado(a) y divorciado(a).

**Proveedor:**

1. Si.
2. No.

Grupo de Ocupación del Empleo:

1. Grupo 1 (Directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleado de oficina) Se agruparon los códigos 1, 2, 3 y 4, según la Clasificación Internacional Uniforme de Ocupaciones (CIUO) 88 a 1 dígito.
2. Grupo 2 (Vendedores, agricultores, pescadores, operarios de construcción, trabajadores de artes, alimentos y otros). Se agruparon los códigos 5, 6, 7 y 8, según la Clasificación Internacional Uniforme de Ocupaciones (CIUO) 88 a 1 dígito.
3. Grupo 3 (Asistentes, peones, ayudantes, vendedores ambulantes, recolectores, otros no identificados). Se agruparon los códigos 9 y 10, según la Clasificación Internacional Uniforme de Ocupaciones (CIUO) 88 a 1 dígito.

Rama de Actividad Económica del Empleo (es el sector al que se dedica la empresa en la que trabajan los ocupado o trabajaban los desocupados anteriormente):

1. Actividad primaria. Se agruparon los códigos 1 y 2, según la Clasificación Industrial Internacional Uniforme (CIIU) revisión 3 – 1 dígito.
2. Actividad secundaria. Se agruparon los códigos 3, 4, 5 y 6, según la Clasificación Industrial Internacional Uniforme (CIIU) revisión 3 – 1 dígito.
3. Actividad terciaria. Se agruparon los códigos 7 al 17, según la Clasificación Industrial Internacional Uniforme (CIIU) revisión 3 – 1 dígito.

Clasificación Internacional de Categorías del Empleo (abreviada CISE) :

1. Independiente. Se agruparon los códigos 1 y 2, según la Clasificación Internacional de la Situación en el Empleo (CISE) 1993.
2. Dependiente. Se agruparon los códigos 3, 4, 5 y 6, según la Clasificación Internacional de la Situación en el Empleo (CISE) 1993.



Capítulo 5

Resultados.

5.1. Análisis regional descriptivo.

Se realizó una comparación de las bases de datos de los meses centrales de Agosto y Noviembre de 2015, no se utilizó el factor de expansión para una “ampliación” de los registros por individuo encuestado, es decir, los datos están en bruto y solo un individuo por registro. Además, se destaca que la categoría ocupados, con respecto a su frecuencia relativa o valor porcentual, carece de relación con la “*Tasa de Ocupación*”, dado que no se incluyó los Inactivos en el análisis. A continuación se observa en la **Tabla 5.1**, un breve resumen general de las 2 bases de datos.

La cantidad de registros o personas de interés encuestadas son de 5231 y 5165 para el mes central de Agosto y Noviembre, respectivamente. Agosto presenta un mayor “*Índice de Desempleo*” (7,9%), con respecto a Noviembre (6,9%). Según la variable de interés, “*Condición de actividad*”, que separa a los encuestados como “ocupado”, “otros ocupados” y “desocupado”, se tiene que en el mes central de Agosto hay 462 personas desocupadas y, para mes central de Noviembre hay 420 personas desocupadas. El porcentaje o tasa bruta de desocupados (proporción entre desocupados y total de la muestra de interés) es de 8,832% y 8,132% para Agosto y Noviembre, respectivamente. Siendo mayor en el mes central de Agosto, al igual que Índice de Desempleo (la variación de decimales se debe a que no se utilizó el factor de expansión).

	Mes central (2015)	Jul-Agosto-Sept.(%)	Oct-Noviembre-Dic.(%)
Datos generales	Índice de desempleo	7,9%	6,9%
	Muestra total	5231	5165
Condición de Actividad	Desocupado	462 (8,832)	420 (8,132)
	Otros Ocupados	687 (13,133)	569 (11,016)
	Ocupado	4082 (78,035)	4176 (80,852)

Tabla 5.1: Información general.

Con respecto a las variables independientes, algunas de tipo sociodemográficas como Provincia, Tipo Estrato, Sexo y Estado Conyugal, se observan las principales características en

la **Tabla 5.2**. Las 4 variables presentaron registros de porcentajes similares para cada mes central, comparando sus respectivas categorías. Por ejemplo, al observar los registros de la variable Provincia, se tuvo que los encuestados pertenecientes a Arauco son un 11 % y en Concepción son un 73 % en ambos meses (con variación en decimales). La variable Tipo Estrato, presentó la mayor cantidad de encuestados en ciudades con un 76 % aproximadamente en ambos meses, mientras que, en área rural hubo 8 % de encuestados en ambos meses aproximadamente. La variable Sexo, presentó un 44,2 % de encuestadas mujeres en Agosto y un 43,5 % en Noviembre, los hombres fueron mayoría con porcentaje complementario en cada mes. La variable Estado Conyugal presentó sus mayores registros en los encuestados casado(a) con 42 % en ambos meses, le sigue los encuestados solteros(as) con 32,8 % y 33,7 %, Agosto y Noviembre respectivamente; los registros de conviviente y otro(a) (viudo(a), separado(a)/anulado(a) o divorciado(a)) se encuentran en 13 % y 11 % aproximado respectivamente, en ambos meses.

Variable	Categorías	Mes central (2015)			
		Jul-Ago-Sept.(%)		Oct-Nov-Dic.(%)	
Provincia	Concepción	3853	(73,657)	3779	(73,166)
	Arauco	600	(11,470)	594	(11,500)
	Biobío	778	(14,873)	792	(15,334)
Tipo Estrato	Ciudades	3970	(75,894)	3934	(76,167)
	Restos áreas urbanas	840	(16,058)	826	(15,992)
	Área rural	421	(8,048)	405	(7,841)
Sexo	Hombre	2917	(55,764)	2920	(56,534)
	Mujer	2314	(44,236)	2245	(43,466)
Estado Conyugal	Casado(a)	2219	(42,420)	2201	(42,614)
	Conviviente	717	(13,707)	645	(12,488)
	Soltero(a)	1714	(32,766)	1740	(33,688)
	Otro(a)	581	(11,107)	579	(11,210)

Tabla 5.2: Información descriptiva segregada por mes.

En la **Tabla 5.3** se presenta otro conjunto de variables de interés, de tipo económicas, las cuales poseen similares características porcentuales comparando de un mes a otro. La variable Proveedor tuvo registros alrededor del 50 % en ambas modalidades para los dos meses. La variable Grupo de Ocupación, la mayoría de los encuestados pertenecen al grupo 2 de: vendedores, agricultores, pescadores, operarios de construcción, trabajadores de artes, alimentos y otros; con un 45 % aproximado en ambos meses. La Rama de Actividad principal de los encuestados es la actividad terciaria, con un 66 % aproximado en ambos meses; solo un 7 %, aproximado de los encuestados pertenece a la actividad primaria, en ambos meses. La variable Clasificación Internacional de Categorías del Empleo (CISE), tuvo en sus registros un 80 % aproximado de encuestados como dependientes, en cada mes central, mientras que, solo un 19 % aproximado están clasificados como independientes. Los totales para cada variable y mes cambiaron con respecto a la muestra total de interés, ya que existe información perdida.

Variable	Categorías	Mes central (2015)			
		Jul-Ago-Sept.(%)		Oct-Nov-Dic.(%)	
Proveedor	No proveedor	2569	(49,111)	2553	(49,429)
	Proveedor	2662	(50,889)	2612	(50,571)
	Total	5231		5165	
Grupo de Ocupación	Grupo 1	1609	(31,002)	1508	(29,482)
	Grupo 2	2327	(44,836)	2305	(45,064)
	Grupo 3	1254	(24,162)	1302	(25,455)
	Total	5190		5115	
Rama de Actividad	Actividad primaria	369	(7,110)	385	(7,527)
	Actividad secundaria	1398	(26,936)	1346	(26,315)
	Actividad terciaria	3423	(65,954)	3384	(66,158)
	Total	5190		5115	
CISE	Independiente	1014	(19,855)	989	(19,600)
	Dependiente	4093	(80,145)	4057	(80,400)
	Total	5107		5046	

Tabla 5.3: Información descriptiva segregada por mes.

La siguiente variable de interés es de tipo discreta y posee 6 modalidades, donde cada valor corresponde al Número de Personas en el Hogar y, la última modalidad corresponde a 6 o más personas que viven en el hogar, el detalle en la **Tabla 5.4** segregado por mes. Como ha sido la lógica de las variables anteriores, en particular de estos registros de trimestres móviles consecutivos, los porcentajes de un mes central a otro varían solo en décimas. Luego, la mayoría de los encuestados vive solo, con un 45% aproximado, le sigue con 2 personas en el hogar el 27,6% aproximado de los encuestados en ambos meses. El porcentaje de registros va disminuyendo a medida que aumenta la cantidad de personas en el hogar, por ejemplo: casi el 3% de los encuestados responde que viven 5 personas en el hogar, para ambos meses; al igual que 6 o más personas en el hogar, son casi el 3% de los encuestados en ambos meses.

Variable	Categorías	Mes central (2015)			
		Jul-Ago-Sept.(%)		Oct-Nov-Dic.(%)	
Número de Personas en el Hogar	1	2406	(45,995)	2349	(45,479)
	2	1447	(27,662)	1427	(27,628)
	3	791	(15,121)	798	(15,450)
	4	314	(6,003)	306	(5,924)
	5	146	(2,791)	138	(2,672)
	6 o más	127	(2,428)	147	(2,846)
	Total	5231	(100,0)	5165	(100,0)

Tabla 5.4: Frecuencia de Número de Personas en el Hogar segregada por mes.

Del grupo de variables independientes, hubo 2 de tipo numéricas que fueron categorizadas en no más de 8 clases, siguiendo recomendaciones de Brigitte (1992). Se observa en la **Tabla**

5.5 la variable Edad y la variable Años de Estudio, la primera fue categorizada en clases con rangos de 10 años cada una, excepto la última clase que corresponde a rangos de jubilación de 65 o más años y, que corresponde a un 5 % de los encuestados para cada mes. La modalidad de mayor frecuencia es de 45 a 54 años con casi un 25 % de los encuestados, le sigue los rangos de edad de 25 a 34 años y de 35 a 44 años, ambos rangos con un 21 % aproximado cada uno, y también en cada mes central. Al comparar los porcentajes de un mes a otro, se observa que son prácticamente idénticos en cada rango etario de comparación.

La variable Años de Estudio, también presenta porcentajes similares en su respectiva modalidad de un mes central a otro. La menor modalidad corresponde a encuestados de 19 o más años de estudio, cerca de un 2 % en ambos meses. La mayoría de los encuestados tiene entre 9 y 12 años de estudio, con un 42 % aproximado en cada mes central; luego un importante 32 % aproximado tiene de 13 a 18 años de estudio en ambos meses y, por último, un 23 % aproximado no tiene estudios o tiene hasta 8 años de estudio, en cada mes central.

Variable	Categorías	Mes central (2015)	
		Jul-Ago-Sept.(%)	Oct-Nov-Dic.(%)
Edad	De 15 a 24 años	568 (10,858)	559 (10,823)
	De 25 a 34 años	1125 (21,506)	1105 (21,394)
	De 35 a 44 años	1104 (21,105)	1060 (20,523)
	De 45 a 54 años	1295 (24,756)	1288 (24,937)
	De 55 a 64 años	877 (16,765)	894 (17,309)
	De 65 o más años	262 (5,009)	259 (5,015)
	Total	5231	5165
Años de Estudio	De 0 a 8 años estudio	1193 (22,833)	1231 (23,871)
	De 9 a 12 años estudio	2208 (42,258)	2188 (42,428)
	De 13 a 18 años estudio	1724 (32,995)	1650 (31,995)
	De 19 o más años estudio	100 (1,914)	88 (1,706)
	Total	5225	5157

Tabla 5.5: Frecuencia de Edad y Años de Estudio segregada por mes.

5.2. Análisis regional mediante Tablas de Contingencias.

Se presentan a continuación resultados bidimensionales, en formato de Tablas de Contingencia, de las 11 variables independientes cruzadas con la Condición de Actividad de modalidades: ocupado, otros ocupados y desocupado, para cada mes central de Agosto y Noviembre. Cada tabla de contingencia presenta su frecuencia absoluta y la frecuencia porcentual por columna, para la Condición de Actividad; además, se observa el estadístico Chi-cuadrado de independencia y su valor p. El objetivo es poder analizar las principales cualidades que se pueden desprender de la Condición de desocupado, por mes central y ver la diferencia entre

meses centrales.

La **Tabla 5.6** muestra las Tablas de Contingencia del grupo de variables sociodemográficas (análogo a la Tabla 5.2) comparadas por mes central.

La variable Provincia para el mes de agosto, presentó el mayor porcentaje de desocupados en Concepción (79%) y el menor en Arauco (7,8%). Por otra parte, los mayores porcentajes para ocupados y otros ocupados también fueron para Concepción, 72,6% y 76,4% respectivamente, aunque el porcentaje de desocupados fue mayor entre las condiciones de actividad.

El mes de Noviembre, presentó un comportamiento similar para los desocupados, la mayoría son de Concepción (77,6%) y los menores registros en Arauco (7,9%). Las otras condiciones de actividad, también presentaron su mayor porcentaje para Concepción. Sin embargo, el porcentaje de otros ocupados en Concepción es mayor (78,2%) que el de desocupados (77,6%) en Concepción.

En ambos meses centrales se observó un valor p significativo ($< 0,05$), por lo que existe relación entre Provincia y Condición de Actividad. Al comparar los meses centrales, se observó que el mes de Agosto está levemente por sobre el mes de Noviembre, con respecto al porcentaje de desocupados que son de Concepción, 79% y 77,6% respectivamente.

La variable Tipo de Estrato para el mes de agosto, presentó el mayor porcentaje de desocupados en ciudades (81,8%) y el menor porcentaje de desocupados es de área rural (4,3%). Las otras condiciones se comportan de igual manera, por ejemplo, el mayor porcentaje de ocupados es en ciudades y el menor en área rural. Entre condiciones, el porcentaje de desocupado de ciudades es mayor que los porcentajes de ocupados que son de ciudades y otros ocupados que son de ciudades (75,5% y 74%, respectivamente).

El mes de noviembre tuvo un comportamiento prácticamente idéntico, el mayor porcentaje de desocupados fue en ciudades (82,9%), siendo también el mayor entre las otras condiciones en ciudades. Los ocupados de ciudades y otros ocupados de ciudades son un 75,4% y 76,6% respectivamente). El menor porcentaje de desocupados fue en área rural, con un 5,9%.

En ambos meses resultó un valor p significativo ($< 0,05$), por lo tanto existe relación entre Tipo Estrato y Condición de Actividad. Se observó un leve mayor porcentaje de desocupados de ciudades en noviembre (82,9%) que los desocupado de ciudades en mes de agosto (81,8%).

La variable Sexo, en el mes de agosto el mayor porcentaje de desocupados son hombre con un 57,4% y los desocupados que son mujer un 42,6%. Para la condición de ocupados el resultado es análogo, en cambio la condición de otros ocupados los resultados son opuestos.

El mes de noviembre tiene resultados equivalentes, el mayor porcentaje de desocupados son hombres con un 58,3%, y un 41,7% de desocupados corresponden a mujeres. Las condiciones

ocupados y otros ocupados también presentan su mayoría en hombres, aunque esta última sólo por décimas porcentuales.

Ambos meses resultaron con un valor p significativo ($< 0,05$), entonces existe relación entre Sexo y Condición de Actividad. Al comparar por mes, el porcentaje de desocupados que son hombres es levemente mayor en Noviembre. *El comportamiento de Sexo en la Condición de ocupado y desocupado es el mismo, hay mayor porcentaje de hombres, algo esperado ya que los encuestados hombres son un 56 % aproximado en ambos meses. Otros ocupados es la Condición que varía su comportamiento en Sexo.*

La variable Estado Conyugal (4 categorías), “en el mes de agosto, mostró que un 47,6 % de los desocupados es soltero(a), le sigue un 31,6 % que es casado(a), el menor porcentaje de desocupados es viudo(a), separado(a) o divorciado(a) con un 6,7 % (categorizado como otro(a)). En cambio, los ocupados y otros ocupados solteros alcanzan un 31,1 % y 32,3 % respectivamente, bastante menor que los desocupados que son solteros. Sucede lo mismo para desocupado que son conviviente (14 %), tienen mayor porcentaje que las otras condiciones que son conviviente. Además, los encuestados ocupados y otros ocupados tienen el mayor porcentaje en la modalidad casado(a) ambos con un 43 % aproximado”.

Mes central	Jul-Agosto-Sept.			Oct-Noviembre-Dic.			
Condición Actividad	Ocupado n=4082 (%)	Otros ocup. n=687 (%)	Desocupado n=462 (%)	Ocupado n=4176 (%)	Otros ocup. n=569 (%)	Desocupado n=420 (%)	
Provincia	χ^2 (valor p)	39,508 (<0,001)			χ^2 (valor p)	17,531 (0,002)	
Concepción	2963 (72,587)	525 (76,419)	365 (79,004)	3008 (72,031)	445 (78,207)	326 (77,619)	
Arauco	461 (11,293)	103 (14,993)	36 (7,792)	501 (11,997)	60 (10,545)	33 (7,857)	
Biobío	658 (16,120)	59 (8,588)	61 (13,203)	667 (15,972)	64 (11,248)	61 (14,524)	
Tipo Estrato	χ^2 (valor p)	14,241 (0,007)			χ^2 (valor p)	16,859 (0,002)	
Ciudades	3083 (75,527)	509 (74,090)	378 (81,818)	3150 (75,431)	436 (76,626)	348 (82,857)	
RAU	653 (15,997)	123 (17,904)	64 (13,853)	678 (16,236)	101 (17,750)	47 (11,190)	
Área rural	346 (8,476)	55 (8,006)	20 (4,326)	348 (8,333)	32 (5,624)	25 (5,952)	
Sexo	χ^2 (valor p)	23,742 (<0,001)			χ^2 (valor p)	10,429 (0,005)	
Hombre	2328 (57,031)	324 (47,162)	265 (57,359)	2389 (57,208)	286 (50,264)	245 (58,333)	
Mujer	1754 (42,969)	363 (52,838)	197 (42,641)	1787 (42,792)	283 (49,736)	175 (41,667)	
Est. Conyugal	χ^2 (valor p)	57,546 (<0,001)			χ^2 (valor p)	114,751 (<0,001)	
Casado(a)	1777 (43,533)	296 (43,086)	146 (31,602)	1844 (44,157)	246 (43,234)	111 (26,429)	
Conviviente	563 (13,792)	89 (12,955)	65 (14,069)	524 (12,548)	77 (13,533)	44 (10,476)	
Soltero(a)	1272 (31,161)	222 (32,314)	220 (47,619)	1339 (32,064)	164 (28,822)	237 (56,429)	
Otro(a)	470 (11,514)	80 (11,645)	31 (6,710)	469 (11,231)	82 (14,411)	28 (6,667)	

Tabla 5.6: Tablas de contingencia de variables sociodemográficas por mes.

En el mes de noviembre, el mayor porcentaje de desocupados es de solteros con un 56,4 %, le sigue casado(a) con un 26,4 %, y también, el menor registro de desocupados es para otro(a) con un 6,6 %. Por otra parte, los ocupados que son solteros son un 32 % y otros ocupados solteros sólo un 28,8 %, por lo tanto los desocupados que son solteros son mayores comparado con las otras condiciones en Estado soltero. Los registros de ocupado y otros ocupados tienen su mayoría en la modalidad casado(a) con un 44 % aproximado, ambos mayor que los

desocupados que son casados(as) (26,4 %).

En ambos meses se observó un valor p significativo ($< 0,05$), o sea, existe relación entre el Estado Conyugal y la Condición de Actividad. Por último, al comparar los desocupados en cada mes central, se observó un aumento de porcentaje de desocupados que son solteros(as) en noviembre (56,4%) con respecto a agosto (47,6%). Los desocupados casado(a) y conviviente de noviembre disminuyeron en relación al mes de agosto.

Se observa en la **Tabla 5.7** las Tablas de Contingencia del grupo de variables económicas cruzadas con la Condición de Actividad, segregada por los meses de interés (análogo a la Tabla 5.3).

La variable Proveedor en el mes de agosto, presentó que el 72% de los desocupados no son proveedor, mientras que, los desocupados que son proveedor el 28% aproximado. Otros ocupados mantiene la misma mayoría en sus categorías, pero ocupado que no es proveedor representa casi un 45%. En general, los desocupados que no son proveedor son mayores que las otras condiciones que no son proveedor. *Los registros mostraron que Proveedor se comporta de manera equitativa en sus categorías, 50% de individuos proveedor y 50% de individuos no proveedor. Al segregar por Condición de Actividad los desocupados no proveedor muestran un alza considerable, siendo un suceso llamativo, ya que el ocupado tiene su mayoría siendo proveedor.*

Los resultados para el mes de noviembre son completamente equivalentes, pero aumentando los desocupados que no son proveedor a un 74,7%, porcentaje mayor con respecto a Agosto. Ambos meses tienen un valor p significativo ($< 0,05$), o sea, existe relación entre Proveedor y la Condición de Actividad.

La variable Grupo de Ocupación posee en el mes de agosto, el mayor porcentaje de desocupados en el Grupo 2 con un 50% aproximado y el menor porcentaje de desocupados pertenecen al Grupo 1 (22%). Las otras condiciones de actividad también poseen su mayor porcentaje en el Grupo 2.

El mes central de noviembre se comportó de manera equivalente, la mayoría de desocupados pertenecen al Grupo 2 con un 45% aproximado. De igual manera, ocupado y otros ocupados tienen su mayor porcentaje en el grupo 2 (44% y 52% aproximado respectivamente). *En ambos meses, el menor porcentaje de desocupado fue el Grupo 1, correspondiente a directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleados de oficina. Luego, los registros generales mostraron que el Grupo 3 es quien tiene menos individuos, observando una variación con respecto a la Condición de desocupado.*

Se observó un valor p significativo ($< 0,05$) en los 2 meses centrales, entonces, existe relación entre Grupo de Ocupación y Condición de Actividad. Comparando los desocupados de ambos meses, agosto tiene un mayor porcentaje de desocupado del Grupo 2, que los del mes

de Noviembre (50 % vs 45 % aproximado). Además, los desocupados de Grupo 1 y Grupo 3, aumentaron en mes de Noviembre, con respecto al mes central de Agosto.

La variable Rama de Actividad en el mes de agosto, presentó su mayor porcentaje de desocupados en la actividad terciaria con un 52 % aproximado, le sigue la actividad secundaria con un 41 % de desocupados y el menor porcentaje de desocupados fue en la actividad primaria (6,6 %). Las otras condiciones de actividad se comportan de manera idéntica, los mayores porcentajes se encuentran en la actividad terciaria. Los desocupados que pertenecen a la actividad secundaria tienen un mayor porcentaje que, ocupado y otros ocupados de actividad secundaria (ambos 25 % aproximadamente).

Para el mes de noviembre, nuevamente se comporta de manera equivalente al mes de agosto las interpretaciones son análogas. *Sobre los registros generales, la actividad terciaria era quien más individuos poseía (66 % aproximado), y la Condición desocupado tiene registros muy por bajo ese valor, un 52 % y un 58 % aproximado en Agosto y Noviembre respectivamente. Los desocupados de actividad secundaria tienen mayor porcentaje que los registros generales (27 % aproximado), un 41 % en Agosto y un 37 % en Noviembre aproximado.*

Ambos meses, tienen un valor p significativo ($< 0,05$), por lo tanto, existe relación entre Rama de Actividades y Condición de Actividad. Al comparar los desocupados por mes central, se observa que los desocupados de actividad terciaria aumentan en mes de noviembre, a un 57 % aproximado y, los desocupados de actividad secundaria disminuyen a un 37 %.

Mes central 2015	Jul-Agosto-Sept.						Oct-Noviembre-Dic.							
Condición Activ.	Ocupado		Otros ocup.		Desocupado		Ocupado		Otros ocup.		Desocupado			
Proveedor	χ^2 (valor p)		154,774 (<0,001)						χ^2 (valor p)		136,173 (<0,001)			
No proveedor	1831	(44,855)	404	(58,806)	334	(72,294)	1922	(46,025)	317	(55,712)	314	(74,762)		
Proveedor	2251	(55,145)	283	(41,194)	128	(27,706)	2254	(53,975)	252	(44,288)	106	(25,238)		
Total	4082	(%)	687	(%)	462	(%)	4176	(%)	569	(%)	420	(%)		
Grupo Ocupación	χ^2 (valor p)		31,923 (<0,001)						χ^2 (valor p)		22,785 (<0,001)			
Grupo 1	1333	(32,656)	183	(26,638)	93	(22,090)	1272	(30,460)	147	(25,835)	89	(24,054)		
Grupo 2	1768	(43,312)	349	(50,801)	210	(49,881)	1838	(44,013)	299	(52,548)	168	(45,405)		
Grupo 3	981	(24,032)	155	(22,562)	118	(28,029)	1066	(25,527)	123	(21,617)	113	(30,541)		
Total	4082	(%)	687	(%)	421	(%)	4176	(%)	569	(%)	370	(%)		
Rama Actividad	χ^2 (valor p)		41,338 (<0,001)						χ^2 (valor p)		17,726 (0,001)			
Actividad primaria	284	(6,957)	57	(8,297)	28	(6,651)	320	(7,663)	46	(8,084)	19	(5,135)		
Actividad secundaria	1052	(25,772)	173	(25,182)	173	(41,093)	1058	(25,335)	151	(26,538)	137	(37,027)		
Actividad terciaria	2746	(67,271)	457	(66,521)	220	(52,257)	2798	(67,002)	372	(65,378)	214	(57,838)		
Total	4082	(%)	687	(%)	421	(%)	4176	(%)	569	(%)	370	(%)		
CISE	χ^2 (valor p)		232,264 (<0,001)						χ^2 (valor p)		163,725 (<0,001)			
Independiente	735	(18,252)	261	(39,605)	18	(4,276)	768	(18,632)	205	(37,004)	16	(4,324)		
Dependiente	3292	(81,748)	398	(60,395)	403	(95,724)	3354	(81,368)	349	(62,996)	354	(95,676)		
Total	4027	(%)	659	(%)	421	(%)	4122	(%)	554	(%)	370	(%)		

Tabla 5.7: Tablas de contingencia de variables económicas por mes.

La variable Clasificación Internacional de Categorías del Empleo (CISE) tiene prácticamente idénticos resultados en ambos meses. El mes de agosto los desocupados en su

mayoría son dependientes (95,7 %) y los independientes son apenas un 4,3 %. Los ocupados y otros ocupados, también tienen su mayoría en sus registros dependientes (81 % y 60 % aproximado respectivamente), sin embargo, la mayoría corresponde a desocupados dependiente. En noviembre, las interpretaciones son análogas.

Los registros generales daban cuenta, efectivamente, que los individuos con Empleo dependiente son mayoría con 80 % aproximado en ambos meses. Sin embargo, en la Condición de desocupado se observan porcentaje de dependiente muy por sobre el registro general, llegan a un 95 % aproximado los desocupados con Empleo dependiente.

Se observó un valor p significativo ($< 0,05$) en ambos meses, entonces hay relación entre CISE y Condición de Actividad. Al comparar ambos meses, los desocupados dependientes se mantienen prácticamente iguales (96 % aproximado ambos).

La variable de tipo discreta con 6 modalidades, **Número de Personas en el Hogar**, se presenta en la **Tabla 5.8** (análoga a Tabla 5.4). En ambos meses posee un valor p significativo ($< 0,05$), por lo tanto, existe relación en el Número de Personas en el Hogar y Condición de Actividad.

En mes de agosto, la mayoría de desocupados viven solos (30 % aproximado), le siguen desocupados donde viven 2 personas en el hogar con casi un 28 % y, desocupados donde viven 3 personas en el hogar con un 27 % aproximado. Mientras que, el menor porcentaje de desocupados son los hogares donde viven 6 o más personas con un 2,4 % aproximado. Las otras condiciones tienen un orden equivalente en sus clases, pero los ocupados y otros ocupados que viven solos tienen un mayor porcentaje (48 % y 40 % aproximado respectivamente) que los desocupados que viven solos.

El mes de noviembre se comporta de forma idéntica y posee análogas interpretaciones. Se destaca que, los desocupados que viven solos disminuyen a un 26 % aproximado en el mes de noviembre y, los desocupados en hogares de 3 personas aumentaron a un 28,5 %.

Mes central	Jul-Agosto-Sept.						Oct-Noviembre-Dic.					
Nº Personas en el Hogar	Ocupado n=4082 (%)		Otros ocup. n=687 (%)		Desocupado n=462 (%)		Ocupado n=4176 (%)		Otros ocup. n=569 (%)		Desocupado n=420 (%)	
1	1990	(48,751)	278	(40,466)	138	(29,870)	2005	(48,012)	235	(41,301)	109	(25,952)
2	1096	(26,850)	223	(32,460)	128	(27,706)	1134	(27,155)	184	(32,337)	109	(25,952)
3	564	(13,817)	103	(14,993)	124	(26,840)	593	(14,200)	85	(14,938)	120	(28,571)
4	236	(5,781)	44	(6,405)	34	(7,359)	231	(5,532)	33	(5,800)	42	(10,000)
5	98	(2,401)	21	(3,057)	27	(5,844)	102	(2,443)	15	(2,636)	21	(5,000)
6 o más	98	(2,401)	18	(2,620)	11	(2,381)	111	(2,658)	17	(2,988)	19	(4,524)
χ^2 (valor p)	110,558 (<0,001)						126,680 (<0,001)					

Tabla 5.8: Tablas de contingencia de personas en el hogar por mes.

Finalmente, la **Tabla 5.9** muestra las variables continuas codificadas (análogas a Tabla 5.5) en relación con la Condición de Actividad.

La variable Edad en el mes de agosto, posee un 25,8 % de desocupados entre 25 y 34 años, le sigue con un 24,7 % los desocupados de 15 a 24 años (habilitados y dispuestos a trabajar, según la definición de desocupados); los menores porcentajes de desocupados son entre 55 y 64 años con un 10,6 % y, desocupados de 65 o más años con un 2,3 %. Las otras condiciones de actividad, poseen distinta estructura o distribución de porcentaje en las edades, sin embargo los menores porcentajes coinciden y se encuentran en edad de jubilación, por ejemplo, los ocupados de 65 o más años son un 4,6 % aproximado. Mientras que, otros ocupados de 65 o más años aumentan un 9,2 % aproximado.

El mes de noviembre se comportó prácticamente igual en las condiciones de ocupado y otros ocupados, también coincide los menores porcentajes en edad de jubilación. Por otra parte, los mayores desocupados se encuentran entre 15 y 24 años con un 30,4 %, le sigue con un 25,7 % los desocupados de 25 a 34 años; los menores porcentajes de desocupados igual son de 55 a 64 y de 65 o más años, con un 8,6 % y 2,9 % respectivamente.

Claramente, los mayores porcentajes de desocupados se ubican a temprana edad, como era de esperarse, sin embargo el porcentaje en el tramo de Edad 15 a 24 años es importante destacar ya que representa la edad escolar y de formación profesional.

Se observó en ambos meses un valor p significativo ($< 0,05$), por lo que, existe relación entre la Edad y la Condición de Actividad. Al comparar los desocupados de ambos meses, en noviembre aumenta la clase de 15 a 24 años, de 24,7 % a un 30,5 %, y disminuyen los desocupados de 45 a 54 años, de 19 % en agosto a un 14,7 % en noviembre.

Mes central 2015	Jul-Agosto-Sept.					Oct-Noviembre-Dic.						
Condición Activ.	Ocupado		Otros ocup.		Desocupado	Ocupado		Otros ocup.		Desocupado		
Edad	χ^2 (valor p)		162,029 (<0,001)			χ^2 (valor p)		235,625 (<0,001)				
De 15 a 24 años	363	(8,893)	91	(13,246)	114	(24,675)	367	(8,788)	64	(11,248)	128	(30,476)
De 25 a 34 años	873	(21,387)	133	(19,360)	119	(25,758)	888	(21,264)	109	(19,156)	108	(25,714)
De 35 a 44 años	891	(21,828)	132	(19,214)	81	(17,532)	886	(21,216)	100	(17,575)	74	(17,619)
De 45 a 54 años	1051	(25,747)	156	(22,707)	88	(19,048)	1094	(26,197)	132	(23,199)	62	(14,762)
De 55 a 64 años	716	(17,540)	112	(16,303)	49	(10,606)	741	(17,744)	117	(20,562)	36	(8,571)
De 65 o más años	188	(4,606)	63	(9,710)	11	(2,381)	200	(4,789)	47	(8,260)	12	(2,857)
Total	4082	(%)	687	(%)	462	(%)	4176	(%)	569	(%)	420	(%)
Años de Estudio	χ^2 (valor p)		54,512 (<0,001)			χ^2 (valor p)		24,393 (0,002)				
De 0 a 8 años estudio	897	(22,001)	220	(32,070)	76	(16,450)	1001	(24,011)	155	(27,289)	75	(17,857)
De 9 a 12 años estudio	1729	(42,409)	247	(36,006)	232	(50,216)	1728	(41,449)	255	(44,894)	205	(48,810)
De 13 a 18 años estudio	1364	(33,456)	212	(30,904)	148	(32,035)	1363	(32,694)	150	(26,408)	137	(32,619)
De 19 o más años estudio	87	(2,134)	7	(1,020)	6	(1,299)	77	(1,847)	8	(1,408)	3	(0,714)
Total	4077	(%)	686	(%)	462	(%)	4169	(%)	568	(%)	420	(%)

Tabla 5.9: Tablas de contingencia de variables categorizadas por mes.

La variable **Años de Estudio** codificada en clases, muestra que en el mes de agosto el 50 % de los desocupados tiene de 9 a 12 años de estudio, el 32 % de desocupados tiene de 13 a 18 años de estudio, le sigue con un 16,5 % desocupados de 0 a 8 años de estudio; el menor porcentaje de desocupado tienen 19 o más años de estudio (1,3 %). Las condiciones de ocupado y otros ocupados tienen idéntica estructura en los registros de los Años de Estudio,

su mayoría de registros se encuentran en el rango de 9 a 12 años de estudio y, la menor cantidad de registros en encuestados de 19 o más años de estudio.

El mes de noviembre no tiene cambios radicales en los porcentajes de desocupados, la mayor cantidad de desocupados proviene de 9 a 12 años de estudio (48,8%) y la menor cantidad es de 19 o más años de estudio (0,7%). Por otra parte, las condiciones de actividad restantes mantienen sus altos y bajos porcentajes en los mismos rangos de Años de Estudio.

“Hay un alto porcentaje de desocupados de 9 a 12 años de estudio (50%), puede ser respuesta a las personas que no terminaron su enseñanza media o que solo tengan enseñanza media completa, este grupo de personas no posee estudios completos y puede que no tenga especialización en enseñanza superior. Le sigue los desocupados de 13 a 18 años de estudio (32%), aquí puede coincidir con lo anterior, jóvenes con poca experiencia o que aún estén estudiando y las empresas no requieran de estudiantes o sin especialización aún”.

Los registros generales mostraron un 42% aproximado de individuos de 9 a 12 años de estudio en ambos meses, sobre este caso, al segregar por Condición de Actividad, los desocupados en ese rango de estudio superan el valor general de encuestados por varios puntos. Llamativamente, los registros generales de 0 a 8 años de estudio estaban alrededor de 23% en ambos meses, la Condición de otros ocupados es quien supera estos registros generales con valores de 32% y 27% aproximado en los meses respectivos.

Ambos meses tienen un valor p significativo ($< 0,05$), entonces existe relación entre los Años de Estudio y la Condición de Actividad.

La comparación de un mes a otro, solo lleva a una leve disminución porcentual en los desocupados de 9 a 12 años de estudio del mes de agosto (50,2%) al mes de noviembre (48,8%), los otros rangos se mantienen casi invariantes.

5.3. Análisis regional mediante ACS

Se presentan los resultados de Análisis de Correspondencias simple (ACS), para cada una de las 11 variables independientes con la variable de interés “Condición de Actividad”, y por cada mes central en la Región del Biobío. El objetivo es realizar una comparación en los meses centrales sobre el comportamiento que sigue la Condición de Actividad, en particular la categoría o modalidad de desocupado, con cada variable en estudio. Se observó y analizó el mapa de modalidades conjunto para cada ACS, junto a sus perfiles y contribuciones. Con respecto a la inercia explicada en cada gráfico, dado que la Condición de Actividad posee 3 modalidades y la cantidad máxima de dimensiones es el mínimo de las categorías menos uno de cada variable en estudio, entonces, si la variable independiente posee 3 o más modalidades, el gráfico con 2 dimensiones estará explicando el 100% de la inercia de las variables categóricas, o sea, la Calidad de representación de la proyección será de un 100% para todos los ACS. Además, si la variable independiente posee 2 modalidades, no es ne-

cesario un ACS y basta con una comparación de sus perfiles [21]. Por lo tanto, siempre se obtendrá una explicación completa de la inercia en cada ACS. De manera similar, al tener como máximo 2 dimensiones, las modalidades estarán representadas en un 100 % por los 2 ejes, así, la Calidad de representación de una modalidad en los ejes será siempre de un 100 %.

Se realizó la comparación de los resultados del ACS del mes central de agosto con el mes central de noviembre, para la misma variable independiente con la Condición de Actividad. Se comenzó con las variables de tipo socio demográfico (análogo a Tabla 5.6), luego con las variables de tipo económica (análogo a Tabla 5.7) y, se terminó la comparación con la variable discreta (análogo a Tabla 5.8) y las variables numéricas codificadas (análogo a Tabla 5.9).

La variable Provincia con Condición de Actividad para el mes central de Agosto, posee una Calidad de la Representación de la proyección en la primera dimensión que explica un 77,5 % de la inercia y, la dimensión 2 explica un 22,5 % de la inercia total.

Con respecto a la Contribución de las modalidades a la inercia del factor o eje, para Provincia, Bío Bío es quien más contribuye a la inercia del factor 1 (72,8 %) y Arauco quien más aporta a la inercia del factor 2 (65 %). Por otra parte, para la Condición de Actividad, la modalidad que más contribuye a la inercia del factor 1 es otros ocupados (86 %) y desocupado a la inercia del factor 2 (91 %).

Con respecto a la Calidad de representación de las modalidades en los ejes, para Provincia, el factor 1 representa mejor a Arauco y Bío Bío, mientras que, el factor 2 representa mejor a Concepción (64 %). Por otra parte, para la Condición de Actividad, el factor 1 representa mejor a ocupado y otros ocupados, y, el factor 2 representa principalmente a desocupado (99,9 %).

En el diagrama de modalidades conjuntas, **Figura 5.1**, se observa que el punto más alejado del perfil medio es desocupado y otros ocupados, por el contrario, las modalidades más similares al perfil medio es Concepción y ocupado. La dimensión 1 opone a otros ocupados en lado negativo de las otras condiciones y, opone a Bío Bío en lado positivo del resto de las Provincias. La dimensión 2 opone por bajo a desocupado del resto de condiciones y, también a Concepción de las otras Provincias. Entonces, bajo el factor 2 se asocia la Condición de desocupado con la Provincia de Concepción.

Además, el factor 1 (que explica un 77,5 %) asocia en lado positivo a ocupado con la Provincia de Bío Bío y, en lado negativo a otros ocupados con la Provincia de Arauco, ambas sobre el eje.

En el mes de Noviembre para las mismas variables y con respecto a la Calidad de la Representación de la proyección, la dimensión 1 explica una 81,9 % de la inercia y la dimensión 2 el resto, 18,1 %.

Con respecto a la Contribución de las modalidades a la inercia del factor o eje, la provincia

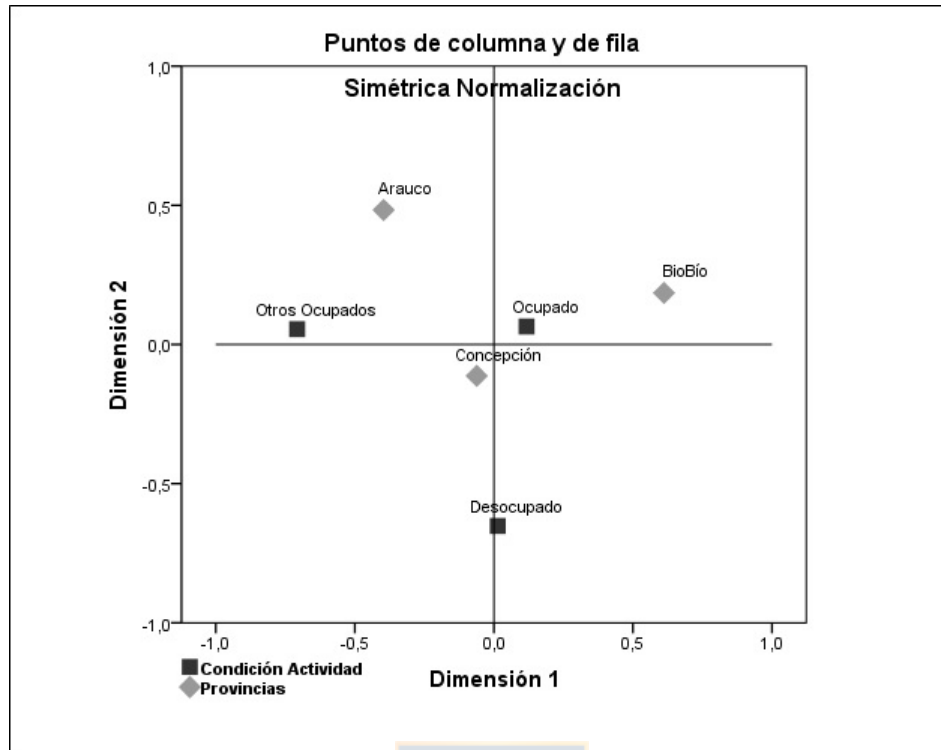


Figura 5.1: ACS de Provincia y Condición de Actividad, Agosto.

que más contribuye es Bío Bío en el factor 1 (42,9%) y de forma muy pareja el resto) y Arauco en el factor 2 (58%), Concepción no contribuye en éste. Por otra parte, las condiciones que más aportan son otros ocupados en el factor 1 (51,9%) y desocupado en el factor 2 (62,3%).

Con respecto a la Calidad de representación de las modalidades en los ejes, las 3 provincias están siendo mayormente representadas por el factor 1 (sobre un 70%). Por otra parte, las 3 condiciones de actividad el factor 1 es quien mejor las representa (sobre un 67%).

En la **Figura 5.2**, se observa que coinciden las modalidades más similares a los perfiles medios, Concepción y ocupado, y el resto de las modalidades se encuentran más alejadas del perfil medio. La dimensión 1 en este caso opone en lado positivo a ocupado del resto de las condiciones y, opone en lado negativo a Concepción de las otras provincias. La dimensión 2 nuevamente opone a desocupado de las otras condiciones. Entonces, en lado negativo se asocia la Condición de desocupado con la Provincia de Concepción y por sobre el eje, que se opone con otros ocupados.

En resumen, no existen cambios sustanciales en los análisis del mes de Agosto al mes de Noviembre, la Condición de desocupado se relaciona principalmente con la Provincia de Concepción sin mayor variación. No existe mayor información o estudio certero que mencione que alguna Provincia de la Región posea constantemente mayores tasas de desocupados, según la información empírica desde el comienzo de la NENE las mayores tasas de desempleo entre

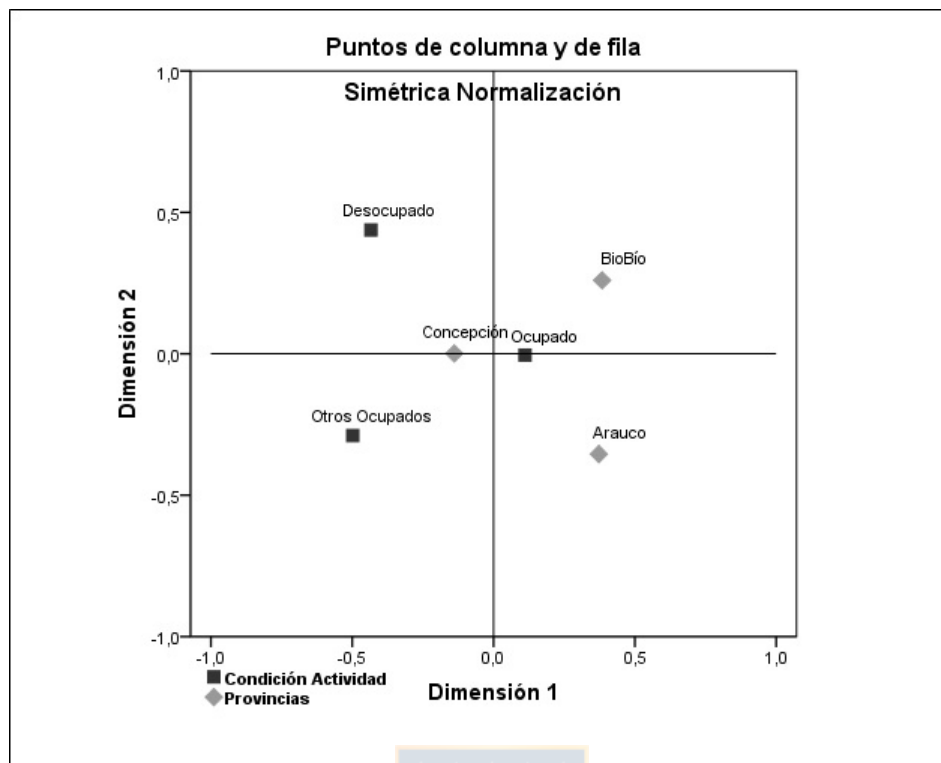


Figura 5.2: ACS de Provincia y Condición de Actividad, Noviembre.

las Provincia de la Región han sido alternadas para Concepción y Bío Bío principalmente [8]. En particular, las tasas de desempleo de las Provincias en los 2 meses centrales en estudio, son mayor en la Provincia de Concepción que las otras provincias.

La variable Tipo de Estrato con Condición de Actividad, para el mes central de Agosto y con respecto a la Calidad de la Representación de la proyección, la dimensión 1 explica un 88,9% de la inercia y, la dimensión 2 explica un 11,1% de la inercia total.

Sobre la Contribución de las modalidades a la inercia del factor, los estratos que más contribuyen son área rural en el factor 1 (67%) y restos de área urbana en el factor 2 (70%). Además, las condiciones que más contribuyen son desocupado al factor 1 (90%) y otros ocupados al factor 2 (83,5%).

La Calidad de representación de las modalidades en los ejes, para Tipo de Estrato sus 3 modalidades son mayormente representadas por el factor 1 (sobre un 61%); y para la Condición de Actividad, el factor 1 representa mejor a desocupado (99,9%) y el factor 2 representa más a otros ocupados (75,5%).

En la **Figura 5.3** del ACS, se observa que las modalidades más distintas del perfil medio son Condición de desocupado y área rural en Estrato. Por el contrario, las modalidades más similares al perfil medio son ocupado y ciudades (CD). La dimensión 1 ubica en lado ne-

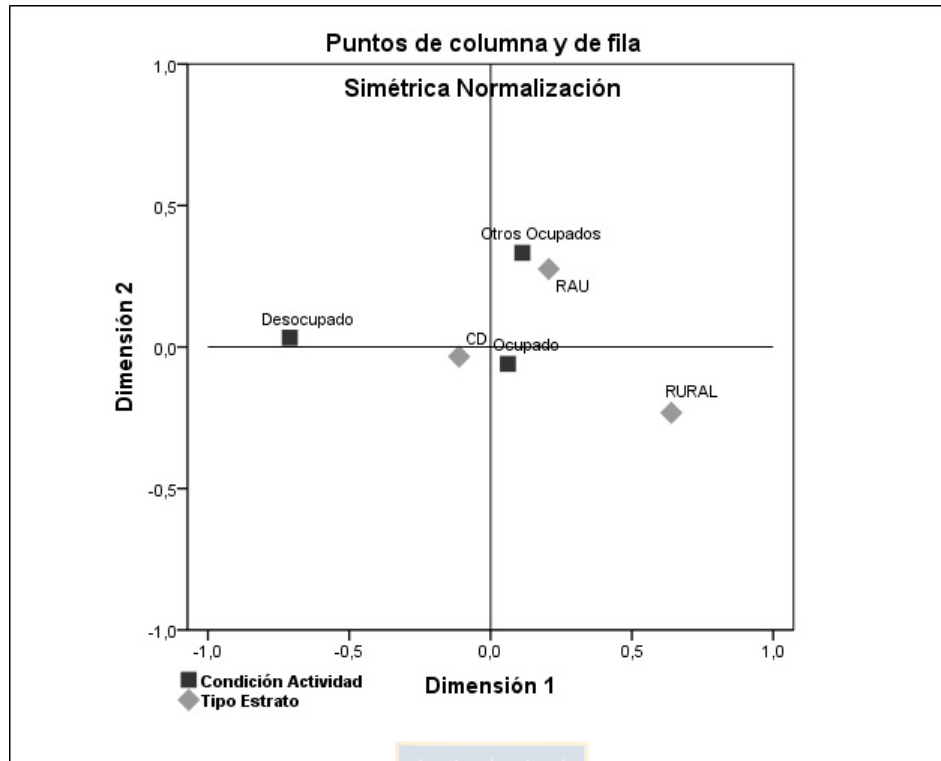


Figura 5.3: ACS de Tipo de Estrato y Condición de Actividad, Agosto.

gativo a desocupado y la opone de las otras condiciones y, también ubica en mismo lado a ciudades oponiéndola de los otros estratos. Por su parte, la dimensión 2 opone levemente a ocupado de las otras condiciones bajo el eje y, opone a RAU de los otros estratos sobre el eje. En conclusión, se asocia la Condición de desocupado con el Estrato ciudades. También, se asocia sobre el factor 2 y a la derecha de la dimensión 1 la Condición de otros ocupados con el Estrato de RAU.

El ACS para el **mes de noviembre** sobre las mismas variables, tiene una Calidad de la Representación de la proyección en la dimensión 1 de un 69,4 % y, la dimensión 2 explica el 30,6 % restante de la inercia.

Con respecto a la Contribución de las modalidades a la inercia del factor, para el Tipo de Estrato, RAU contribuye más a la inercia del factor 1 (45,5 %) y área rural quien más lo hace a la inercia del factor 2 (61,4 %). Además, para la Condición de Actividad, desocupado contribuye más al factor 1 (87 %) y otros ocupados es quien más lo hace al factor 2 (87,6 %).

Sobre la Calidad de representación de las modalidades en los ejes, los 3 estratos están mejor representados por el factor 1 (sobre un 53 %). Por otra parte, para la Condición de Actividad, el factor 1 representa mejor a desocupado (97,6 %) y, el factor 2 representa mejor a otros ocupados (96,6 %).

En la **Figura 5.4** se observa que las modalidades más similares al perfil medio son ocupado

y ciudades, lo que coincide con el mes de agosto. Por el contrario, los puntos más diferentes del perfil medio son desocupado y rural, también se aleja en este mes otros ocupados. La dimensión 1 opone a ocupado de las otras condiciones en lado positivo y, ubica en lado negativo a ciudades que la opone del resto de estratos. A su vez, la dimensión 2 separa por sobre el eje a otros ocupados y RAU del resto de las modalidades. En conclusión, a la izquierda de la dimensión 1 y bajo el eje se asocia la Condición de desocupado con el Estrato de ciudades (CD). Se reitera la relación de otros ocupados con RAU y de ocupado con Estrato rural.

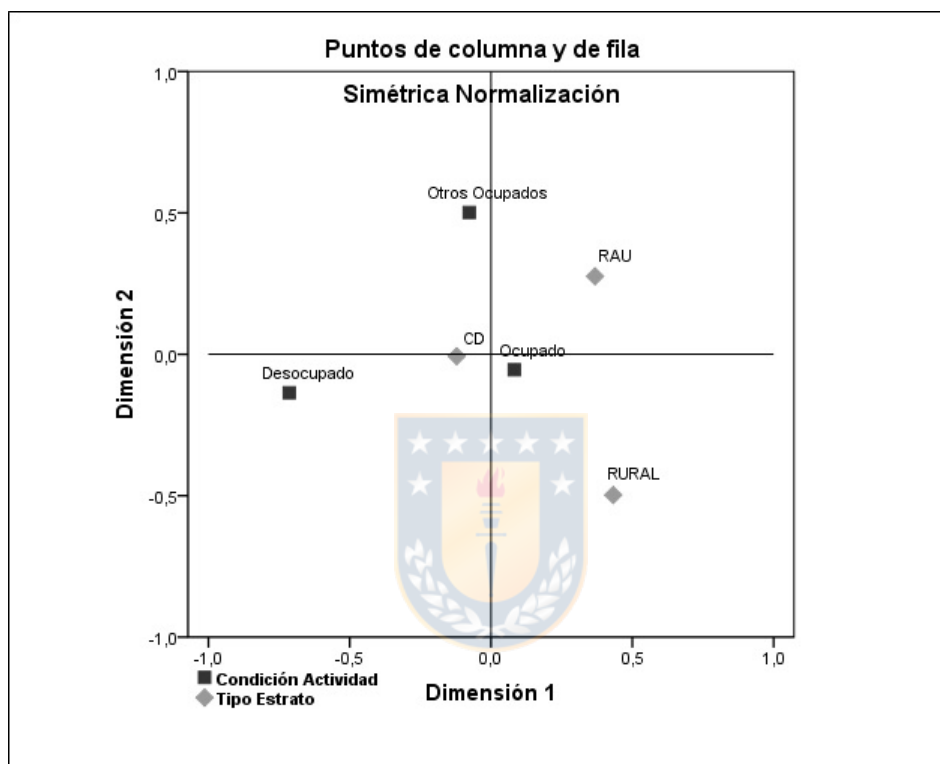


Figura 5.4: ACS de Tipo de Estrato y Condición de Actividad, Noviembre.

Las relaciones encontradas en Tipo de Estrato con Condición de Actividad no sufren variación de un mes central a otro. En particular, la relación de desocupado con ciudades (CD, o grandes centros urbanos) es evidente en ambos análisis. Los gráficos son casi equivalentes, se observa principalmente una leve rotación de los ejes. Este resultado es algo esperado, teniendo en cuenta que una gran concentración de encuestados de Estrato ciudades (90% aproximado) provienen de la Provincia de Concepción, la cual en análisis anterior se encontró una relación con la Condición de desocupado.

La variable Sexo de 2 modalidades, cruzada con Condición de Actividad forma una tabla de 2x3, como se observó en **Tabla 5.6**. Por lo tanto, la dimensión máxima es una y no hay necesidad de llevar a cabo un ACS. Podemos expresar todos los resultados en porcentaje, es más, podemos reinterpretar la tabla de contingencia. Por ejemplo, para el mes de agosto los desocupados y ocupados son mayormente hombres (57% aproximado), mientras que, para la

Condición de otros ocupados el mayor registro es para la categoría mujer (53 % aproximado). Se destaca que el perfil de desocupado y el de ocupado son similares, ambos poseen mayor porcentaje de registros hombre. En el mes de noviembre, los resultados son equivalentes aunque otros ocupados tiene categorías de Sexo casi equivalentes, 50,3 % hombre y 49,7 % mujer.

Por lo tanto, en ambos meses la Condición de desocupado está más relacionado al Sexo hombre, al igual que la Condición de ocupado, luego ambos perfiles son similares. Por el contrario, la Condición otros ocupados está más relacionado al Sexo mujer. Como es sabido, históricamente la mujer es quien siempre posee mayor tasa de desocupación en el país y, según información empírica sucede lo mismo en la Región [8], inclusive también lo confirma Astete B. [9], salvo algunos meses de ciertos años. Los resultados obtenidos para mes de noviembre concuerdan con las tasas de desocupación reales de la Región del Biobío, ya que el mes de noviembre de 2015 la Región presentó una mayor tasa de desocupación en hombre que la tasa de desocupación de mujer. Para mes de Agosto de 2015 en la Región, la tasa de desocupación en mujer fue mayor por 5 décimas porcentuales que la tasa de desocupación de hombre. Estos cambios y variaciones se deben principalmente al trabajo con la base en bruto, sin el factor de expansión.

La variable Estado Conyugal con Condición de Actividad para el mes central de Agosto tiene una Calidad de la Representación de la proyección en la dimensión 1 que explica un 99,3 % de la inercia total de la tabla y la dimensión 2 los restantes decimales.

Con respecto a la Contribución de las modalidades a la inercia del factor, el Estado Conyugal que más aporta a la inercia es, soltero en la dimensión 1 (60 %) y conviviente en la dimensión 2 (76,5 %). Por otro lado, la Condición de Actividad que más aporta a la inercia de cada dimensión es, desocupado en la dimensión 1 (90,9 %) y otros ocupados en la dimensión 2 (86,5 %).

Con respecto a la Calidad de representación de las modalidades en los ejes, todas las modalidades están representadas en el factor 1 (sobre 98 %) excepto el Estado Conyugal de conviviente que está mejor representada en el factor 2 (89,6 %) y, la Condición de otros ocupados que también está mejor representada en el factor 2 (61 %).

La **Figura 5.5** muestra los resultados del ACS, se observa que el punto más similar al perfil medio es ocupado y, el punto más diferente del perfil medio es desocupado. La dimensión 1 opone claramente en lado positivo a desocupado con las otras condiciones, mientras que respecto a Estado Conyugal, en lado derecho soltero y conviviente (cerca del perfil medio) se oponen del resto de modalidades. La dimensión 2 opone levemente algunas modalidades, aunque con poca representatividad por la baja inercia explicada. En conclusión, en lado positivo se asocia principalmente a la Condición de desocupado con el Estado soltero(a).

El mes de noviembre presenta una Calidad de Representación de la proyección similar al mes anterior, la dimensión 1 explica un 96,3 % de la inercia y la dimensión 2 solo un 3,7 %.

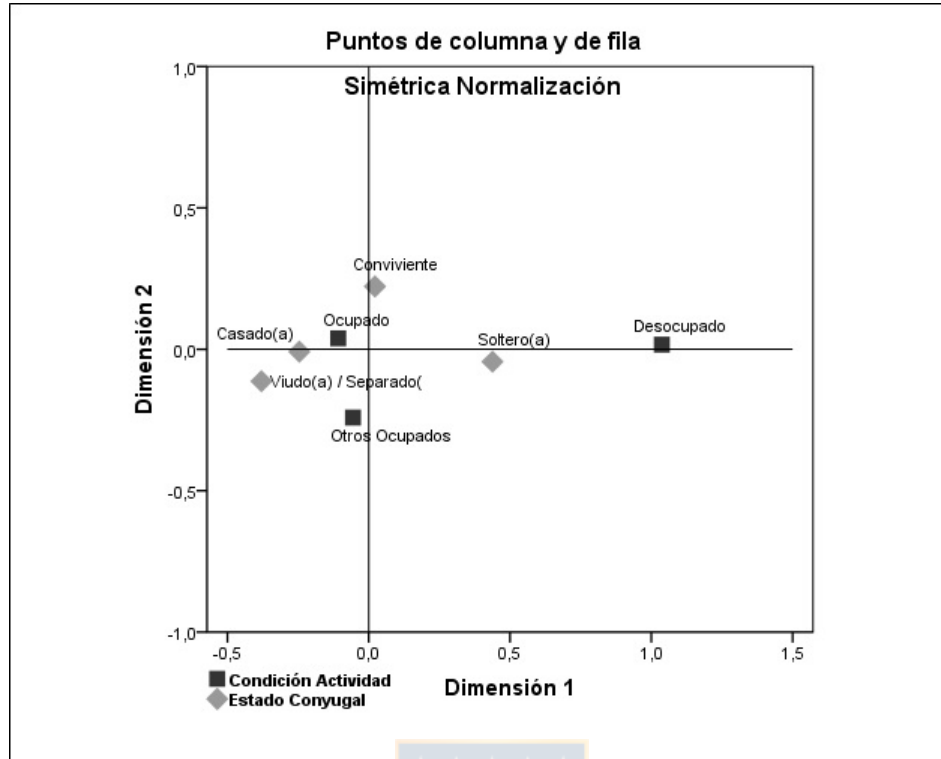


Figura 5.5: ACS de Estado Conyugal y Condición de Actividad, Agosto.

Con respecto a la Contribución de las modalidades a la inercia del factor, el Estado Conyugal que más contribuye a la inercia es soltero(a) en el factor 1 (65 %) y viudo/separado/divorciado(a) en el factor 2 (66 %). Por el contrario, la Condición que más aporta a la inercia es desocupado en el factor 1 (90 %) y otros ocupados en el factor 2 (83 %).

Además, con respecto a la Calidad de representación de las modalidades en los ejes, las 4 modalidades de Estado Conyugal están siendo representadas en su mayoría por el factor 1 (sobre un 78 %). También, es el factor 1 quien mejor representa las 3 modalidades de Condición de Actividad (sobre un 64 %).

Los resultados del ACS (**Figura 5.6**) muestran gran similitud con el mes anterior. Los puntos más similar y diferente del perfil medio coinciden, ocupado y desocupado respectivamente. La dimensión 1 opone completamente en lado derecho la Condición de desocupado y el Estado de soltero(a) del resto de las modalidades. Luego, la dimensión 2 en el lado izquierdo se encarga de separar las condiciones de ocupado y otros ocupados con algunos estados conyugales. En conclusión, se observa claramente que se asocia la Condición de desocupado con el Estado soltero(a) nuevamente.

La situación de Estado Conyugal es bastante clara para ambos meses, ser soltero se relaciona a la Condición de desocupado. Por otra parte, los encuestados casado(a) tienen relación con Condición de ocupado. Los resultados siguen una cierta coherencia, dado que el individuo casado(a) debe tener mayor estabilidad si tiene mayor responsabilidad, por ende una necesi-

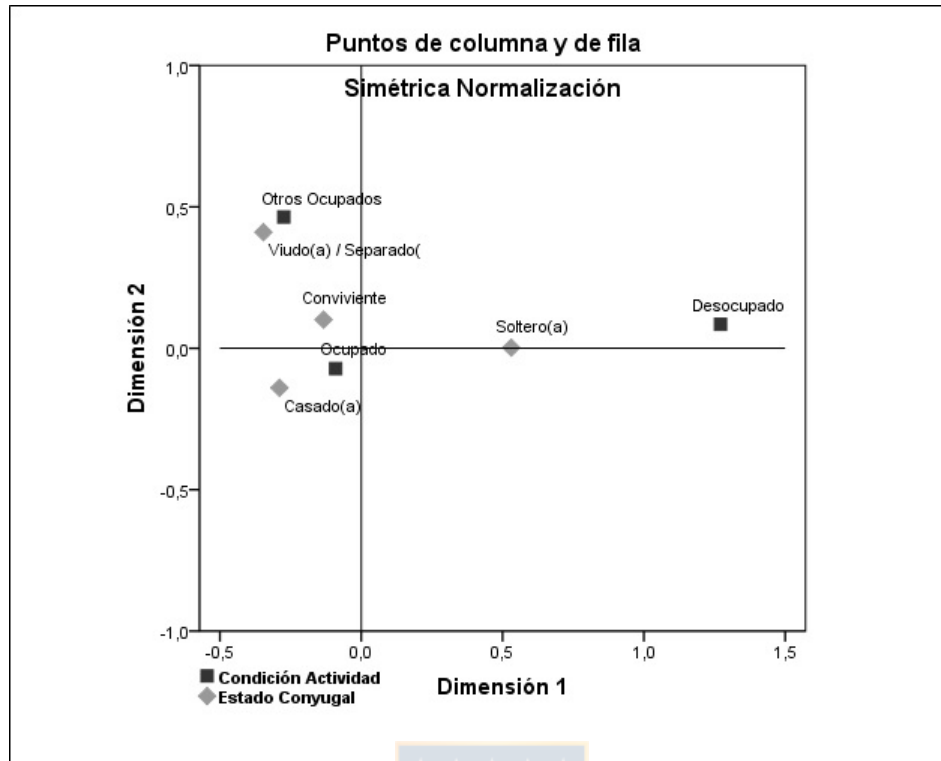


Figura 5.6: ACS de Estado Conyugal y Condición de Actividad, Noviembre.

dad innata de estar empleado. En cambio, el soltero puede ser más dinámico en la búsqueda y elección de empleo, dándose el lujo inclusive de no tener una gran necesidad de estar empleado si se encuentra en un escenario que no vive solo, donde ya se especula que no es jefe de hogar y es joven con poca experiencia, entre otros.

Del grupo de variables económicas, se analiza **Proveedor con Condición de Actividad**. Proveedor posee 2 modalidades, de igual forma que la variable Sexo, por ende, basta con un análisis de los perfiles de interés para poder comparar los mayores porcentajes y encontrar relaciones. Para el mes central de agosto, los desocupados son mayormente no proveedor con un 72,3 %, al igual que otros ocupados. En cambio, para los ocupados se invierte la situación, siendo la mayoría proveedor (55 % aproximado). El mes central de noviembre se comporta de manera equivalente, muestra un amplio porcentaje de desocupados que no son proveedor (74,8 %) y, los ocupados que son proveedor mantienen la mayoría con 54 % aproximado.

En definitiva, la variable Proveedor mantiene la mayoría porcentual y clara relación de la Condición desocupado con no proveedor en ambos meses de estudio. Por el contrario, la Condición de ocupado tiene mayor relación y porcentaje en proveedor. Esto se explica de manera trivial, ya que los individuos que responden con ser proveedores están teniendo una participación activa en el mercado laboral, por ende son individuos ocupados y su responsabilidad es buscar la forma de tener ingresos por trabajo formal o informal. Muy por el contrario, los encuestados que dicen no ser proveedor, pueden no tener una necesidad de recibir ingresos,

dado que otro individuo en el hogar es quien toma esa responsabilidad.

La variable **Grupo de Ocupación con Condición de Actividad**, para el mes central de agosto tiene una Calidad de la Representación de la proyección en la dimensión 1 que explica el 88,2% de la inercia y la dimensión 2 explica el 11,8% restante de la inercia total.

Sobre la Contribución de las modalidades a la inercia del factor, el Grupo de Ocupación que más aporta a la inercia es el grupo 1 en el factor 1 (65,6%) y grupo 3 en el factor 2 (73,8%). Por otro lado, la Condición que más contribuye a la inercia es desocupado en el factor 1 (50,5%) y otros ocupados en el factor 2 (57,5%).

La Calidad de representación de las modalidades en los ejes, para el Grupo de Ocupación el factor 1 representa mejor a grupo 1 y 2 (sobre un 91%) y, el factor 2 representa mejor al grupo 3 (82,8%). Mientras que, la Condición de Actividad tiene sus 3 modalidades representadas principalmente por el factor 1 (sobre un 79%).

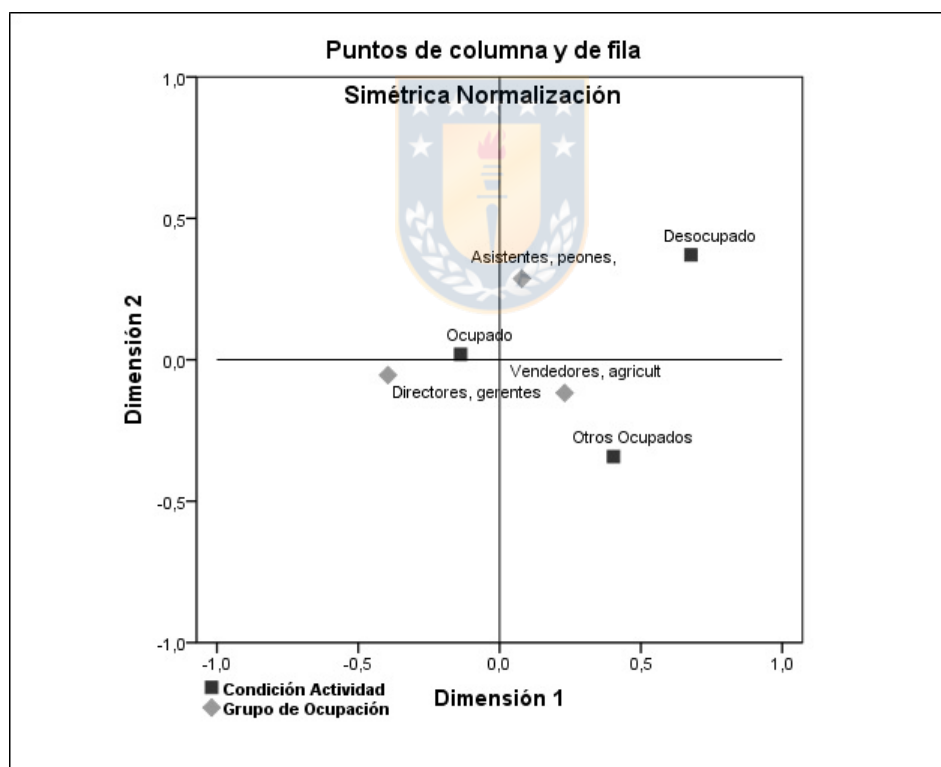


Figura 5.7: ACS de Grupo de Ocupación y Condición de Actividad, Agosto.

Al observar la **Figura 5.7**, de puntos de las modalidades de las variables, la Condición de ocupado es la más similar al perfil medio y la modalidad más diferente del perfil medio es la Condición de desocupado. La dimensión 1 ubica en lado negativo al grupo 1 y la opone de los otros grupos, de igual forma sucede para la Condición de ocupado con las otras condiciones. A su vez, la dimensión 2 separa a grupo 3 sobre el eje del resto de grupos, y separa bajo

el eje a otros ocupados del resto de condiciones. En conclusión, se relaciona principalmente sobre el eje y en lado derecho la Condición de desocupado con el grupo 3 de ocupación que representan a asistentes, peones, ayudantes, vendedores ambulantes, recolectores, otros no identificados. En mismo lado que desocupado se encuentra el grupo 2 de ocupación, pero se encuentra bajo el eje más próximo a otros ocupados.

Además, se observa una clara asociación de la Condición de ocupado con el grupo 1, que corresponde a directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleados de oficina.

En el mes de noviembre para Grupo de Ocupación y Condición de Actividad, el ACS presenta una Calidad de la Representación de la proyección en la dimensión 1 de un 65 % y la dimensión 2 explica el resto de la inercia total, un 35 %.

Sobre la Contribución de las modalidades a la inercia del factor, para el Grupo de Ocupación, el grupo 2 es quien contribuya mayormente a la inercia del factor 1 (54,7 %) y el grupo 3 aporta más a la inercia del factor 2 (58,4 %). Por otro lado, la Condición que más aporta a la inercia es otros ocupados al factor 1 (85,8 %) y desocupado al factor 2 (92 %).

Con respecto a la Calidad de representación de las modalidades en los ejes, el Grupo de Ocupación mejor representado por el factor 1 es el grupo 2 (99,7 %) y por el factor 2 es el grupo 3 (66 %). Mientras que, la Condición de desocupado está mejor representada por el factor 2 (98,5 %) y las otras condiciones por el factor 1 (sobre 83 %).

Se observa de la **Figura 5.8** que la modalidad más similar al perfil medio es la Condición de ocupado y la modalidad más alejada del perfil medio es la Condición de desocupado, que coinciden con el mes de agosto. La dimensión 1 ubica en lado positivo a ocupado que la opone del resto de las condiciones y, ubica en lado negativo a grupo 2 que lo opone de los otros grupos de ocupación. La dimensión 2, que tiene más relevancia en este ACS, ubica bajo el eje la Condición de desocupado y el Grupo de Ocupación 3. Por lo que, se relaciona la Condición de desocupado con el grupo 3 de ocupación. Luego, las otras asociaciones se ubican en la parte superior del factor 2, entre ocupado con grupo 1 de lado derecho y otros ocupados con grupo 2 de ocupación en lado izquierdo.

Las relaciones encontradas en ambos meses de estudio no sufren cambios, es más, claramente cada Condición de Actividad se relacionaba a cada Grupo de Ocupación. Por ejemplo, la Condición de ocupado se relaciona al grupo 1 de Ocupación, siendo estos individuos ocupados en cargos altos, de mayor responsabilidad y compromiso, que requieren una especialización de nivel superior o técnico de nivel medio. Luego, la Condición de desocupado tuvo una principal relación con el grupo 3 de Ocupación, siendo individuos en ocupaciones elementales o trabajadores no calificados como: asistentes, peones, ayudantes, vendedores ambulantes, recolectores, junto a otros individuos no identificados. Aunque el grupo 2 de Ocupación siempre fue ubicado en proximidad con desocupado (por su posición en la representación de la dimensión de mayor inercia), se relacionaba principalmente con la Condición de otros ocupados.

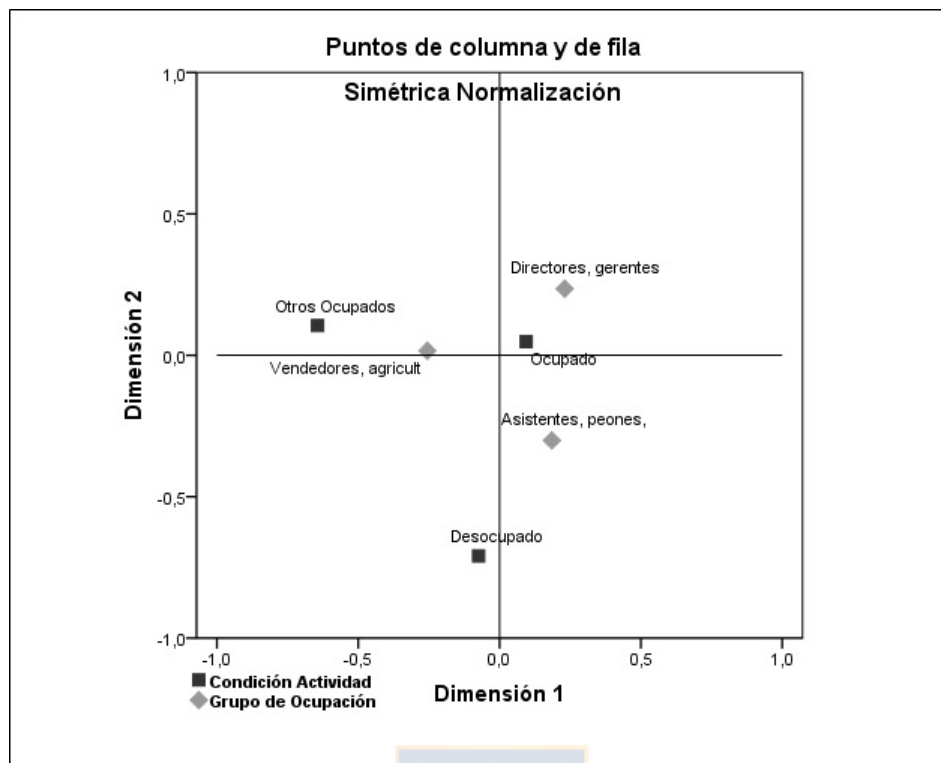


Figura 5.8: ACS de Grupo de Ocupación y Condición de Actividad, Noviembre.

La variable Rama de Actividad económica del empleo con Condición de Actividad, para el mes de agosto presenta una Calidad de la Representación de la proyección en la dimensión 1 del 96,8 % y la dimensión 2 explica el resto de la inercia con un 3,2 %.

Sobre la Contribución de las modalidades a la inercia del factor, la Rama de Actividad que más contribuye a la inercia es actividad secundaria en el factor 1 (72 %) y actividad primaria al factor 2 (92,5 %). De igual manera, la Condición que más aporta a la inercia es desocupado al factor 1 (91,8 %) y otros ocupados al factor 2 (85 %).

La Calidad de representación de las modalidades en los ejes, para Rama de Actividad el factor 2 representa mayormente a actividad primaria (90 %) y el factor 1 a las otras actividades (sobre un 99 %). Luego, para Condición de ocupado y desocupado están siendo representadas principalmente por el factor 1 (sobre 92 %) y el factor 2 representa mejor a otros ocupados (61,4 %).

En la **Figura 5.9** se puede observar que las modalidades más similares al perfil medio son la Condición de ocupado y la actividad terciaria, por otra parte, la modalidad más diferente del perfil medio es la Condición de desocupado. La dimensión 1 ubica en lado negativo la Condición de desocupado y la opone con el resto de condiciones, al igual que actividad secundaria de las otras actividades. La dimensión 2 de poca inercia explicada, sólo separa en lado derecho del factor 1 y sobre el eje a otros ocupados y actividad primaria, a su vez bajo el eje la Condición de ocupado y la actividad terciaria. Por lo tanto, se relaciona en

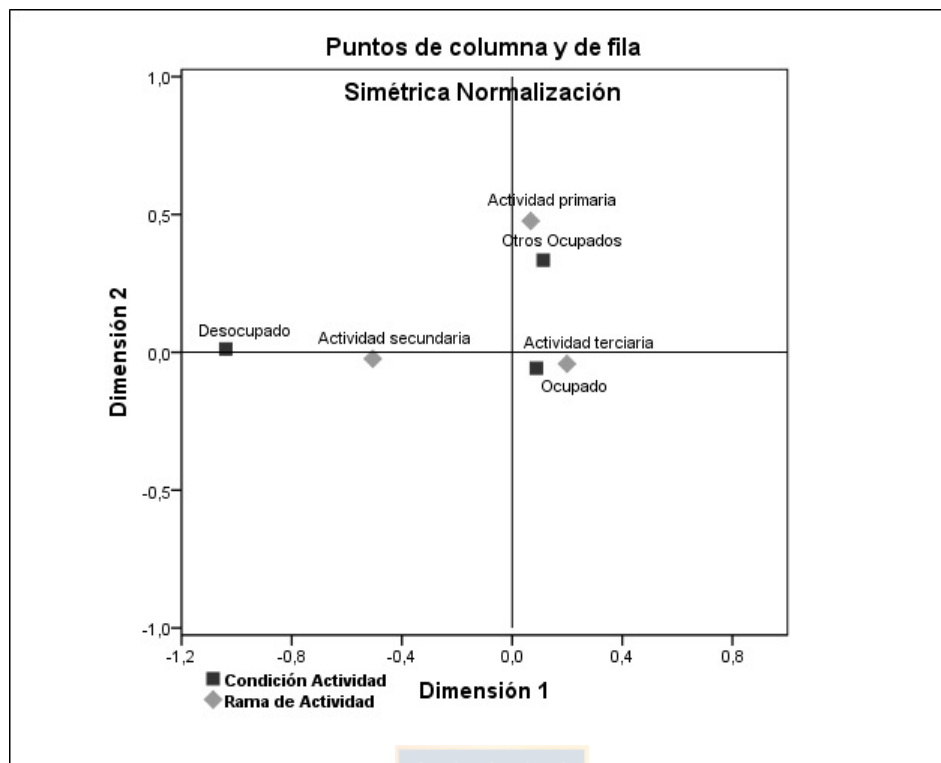


Figura 5.9: ACS de Rama de Actividad y Condición de Actividad, Agosto.

lado izquierdo la Condición de desocupado con la Rama de Actividad secundaria. Además, se asocia otros ocupados con la actividad primaria y ocupado con la actividad terciaria.

El ACS de noviembre para Rama de Actividad, tiene una Calidad de la Representación de la proyección en la dimensión 1 de 98,7% y el resto de la inercia lo explica la dimensión 2.

Con respecto a la Contribución de las modalidades a la inercia del factor, la Rama de Actividad que más aporta a la inercia es la actividad secundaria para el factor 1 (71%) y la actividad primaria en el factor 2 (80,9%). Por otro lado, la Condición de Actividad de desocupado es quien más contribuye a la inercia del factor 1 (91,8%) y otros ocupados aporta más al factor 2 (88,9%).

La Calidad de representación de las modalidades en los ejes para Rama de Actividad, tiene que el factor 1 es el que mejor representa las 3 modalidades (sobre 91%), resultados esperados dada la calidad de representación de la proyección de la dimensión 1. Mientras que, la Condición de ocupado y desocupado están representadas principalmente por el factor 1 (sobre 98%) y, el factor 2 representa mejor a otros ocupados (99,9%).

Al observar la **Figura 5.10**, se concluyó que las modalidades más similares al perfil medio son la Condición de ocupado y la actividad terciaria, por otra parte, el punto o modalidad más distinta del perfil medio es la Condición de desocupado. Luego, la dimensión 1 ubica en

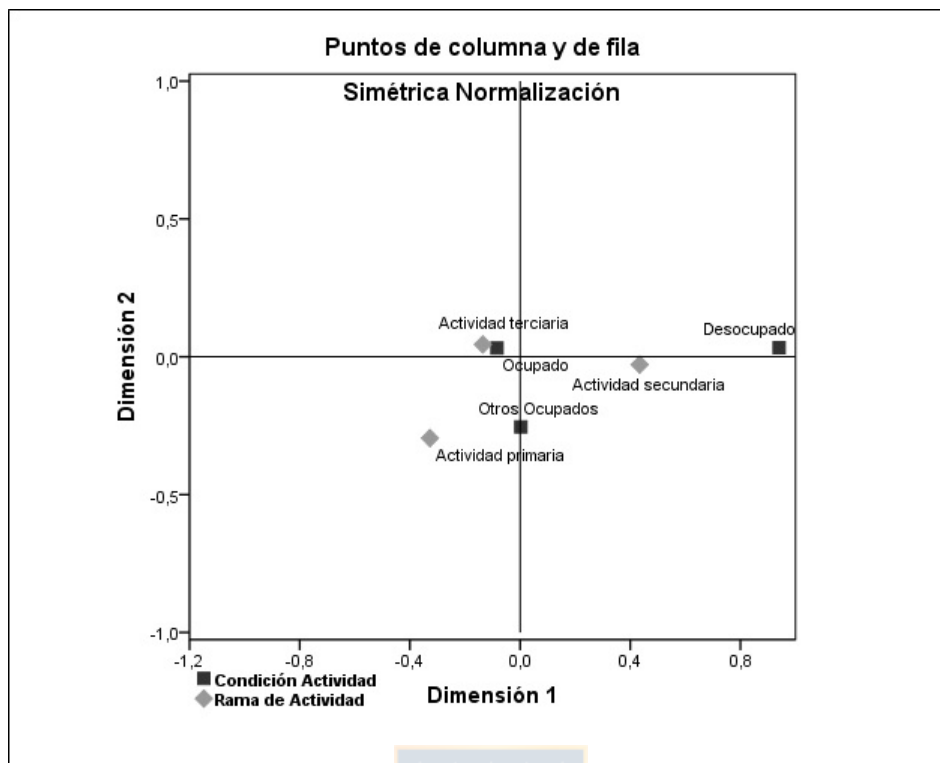


Figura 5.10: ACS de Rama de Actividad y Condición de Actividad, Noviembre.

lado derecho a desocupado y actividad secundaria, y las opone del resto de las modalidades. Además, la dimensión 2 opone levemente sobre el eje la Condición de ocupado y la actividad terciaria de otros ocupados y actividad primaria, principalmente en lado izquierdo de la dimensión 1. Por último, en lado positivo de la dimensión 1 (que mayor representa la inercia) se relaciona la Condición de desocupado y la actividad secundaria.

Al observar cuidadosamente los resultados de ambos ACS, se concluye que las modalidades poseen equivalentes posiciones en el mapa, solo que para un mes las coordenadas están invertidas con respecto al otro mes. La relación existente entre cada modalidad está bien identificada, las condiciones ocupados y otros ocupados se mantienen próximas a actividad terciaria y primaria respectivamente. Mientras que, la Condición de desocupado se relaciona principalmente con la actividad secundaria en ambos meses de estudio. Se destaca que los individuos de actividad secundaria corresponden al sector económico al cual se dedica la empresa de trabajo o en la que trabajaba anteriormente, estas actividades son explotación de minas y canteras, industrias manufactureras, suministro de electricidad, gas y agua, y construcción.

La siguiente variable en estudio es **Clasificación Internacional de Categorías del Empleo (CISE)** la que también posee 2 modalidades, independiente y dependiente. Por lo que, sus resultados son análogos a la tabla de porcentajes interpretada en sección de Tablas de Contingencia (Tabla 5.7).

Con respecto al mes central de agosto, las 3 condiciones de actividad tienen sus mayores porcentajes de registros en la modalidad dependiente, por ejemplo: los desocupados que son dependientes alcanzan un 95,7%, los ocupados dependientes son un 81,7% y, otros ocupados con empleo dependiente son un 60,4%. El porcentaje marginal para CISE corresponde a 19,9% de individuos independiente y a 80,1% para individuos dependiente (ver Tabla 5.3), luego se puede concluir que la Condición de desocupado se encuentra por sobre el perfil medio en la modalidad dependiente, al igual que la Condición de ocupado con empleo dependiente (ambos sobre un 80,1%). Mientras que, la Condición de otros ocupados se encuentra por sobre el perfil medio de empleo independiente, con un 39,6%. Así, principalmente la Condición de desocupado se relaciona más con empleo dependiente y la Condición de ocupado, según su estructura de porcentajes de categoría del empleo, es más similar al perfil medio.

Sobre el mes central de noviembre, los resultados son equivalentes, la Condición de desocupado tiene su mayor porcentaje en empleo dependiente con un 95,7% y está por sobre el perfil columna medio. A su vez, la Condición de ocupado es similar al perfil medio y, la Condición de otros ocupados en empleo independiente se encuentra por sobre el perfil medio con un 37%.

En ambos meses, no hay cambios de relación y sus interpretaciones análogas, en conclusión y por sobre todo la Condición de desocupado se relaciona al empleo dependiente. El trabajador catalogado como independiente, busca por sus medios, por su cuenta e inclusive puede poseer sus propios conocimientos para poder surgir como emprendedor y crear su propio empleo, siendo independiente de cualquier jefe. Este individuo hace lo posible por mantener su condición de ocupado, ya que solo depende de él obtener sus propios beneficios económicos. Por el contrario, el individuo de empleo dependiente, sufre del dinamismo de la economía, los despidos por necesidad de la empresa, la simple reducción de personal, los términos de faenas y en general de la incertidumbre de no saber cuándo pueda ser parte de los desempleados.

A continuación, las variables para ACS son tipo numérica discreta y codificada en clases. **La variable Número de Personas en el Hogar**, para el mes de agosto presenta una Calidad de la Representación de la proyección en la primera dimensión de un 91,3% y, la dimensión 2 representa un 8,7% de la inercia total.

Con respecto a la Contribución de las modalidades a la inercia del factor, para el Número de Personas en el Hogar las modalidades que más aportan a la inercia son 3 y 1 persona en el hogar en el factor 1 (44,5% y 35,5% respectivamente) y la modalidad de 2 personas en el hogar en el factor 2 (63,5%). Por otra parte, la Condición de Actividad que más contribuye a la inercia del factor 1 es desocupado (83,4%) y otros ocupados en el factor 2 (84%).

Además, sobre la Calidad de representación de las modalidades en los ejes, el Número de Personas en el Hogar que están mejor representadas en el factor 1 son las modalidades de 1, 3, 4 y 5 (todas sobre 95%), mientras que, el factor 2 representa mejor 2 y 6 o más personas en el hogar (sobre 91%). Luego, la Condición de Actividad que está mejor representada en

el factor 1 es ocupado y desocupado (sobre 94,8%), por el contrario, otros ocupados está mejor representada en el factor 2 (75 %).

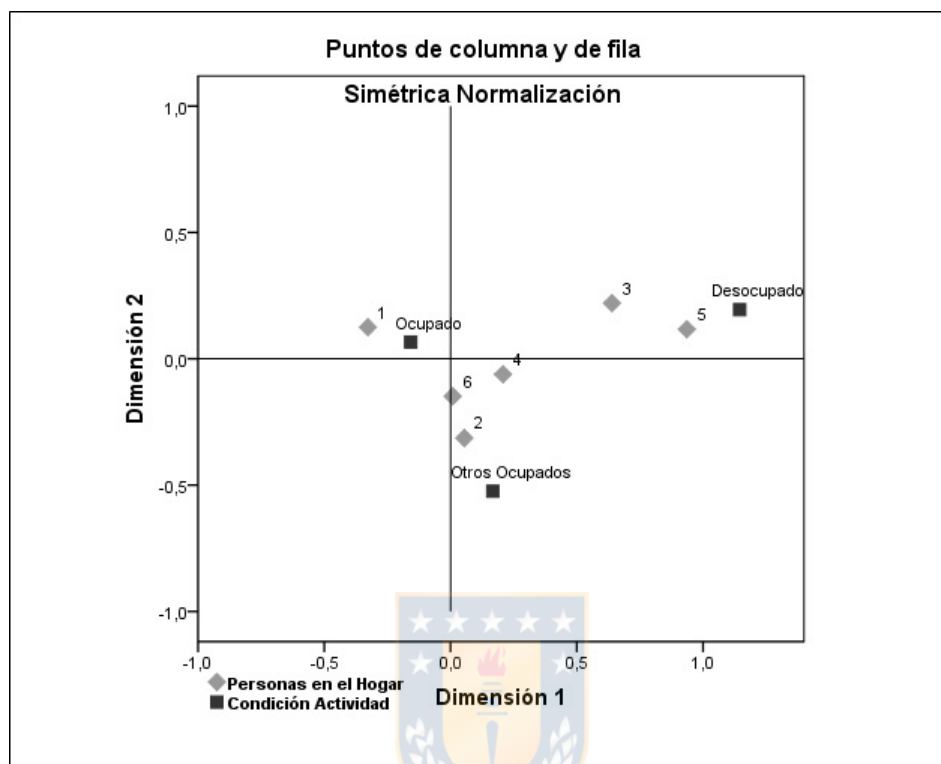


Figura 5.11: ACS de N° de Personas en el Hogar y Condición de Actividad, Agosto.

En la **Figura 5.11**, se observa que las modalidades más cercanas, y por ende similares al perfil medio, son ocupado y 6 o más personas en el hogar. Por otro lado, las modalidades que más difieren del perfil medio son desocupado y 5 personas en el hogar. La dimensión 1 ubica a ocupado en lado izquierdo y la opone del resto de condiciones y, ubica a 1 persona en el hogar en mismo lado y la opone del resto de Número de Personas en el Hogar. La dimensión 2, principalmente separa por sobre y bajo el eje la Condición de desocupado y otros ocupados en lado derecho del gráfico, también separa algunas modalidades del Número de Personas en el Hogar. Por lo que, a la derecha del factor 1 y por sobre el eje se relaciona la Condición de desocupado con 3 y 5 personas en el hogar. Además, en lado izquierdo del factor 1 se asocia la Condición de ocupado con una persona en el hogar.

En el mes central de noviembre, la Calidad de la Representación de la proyección tiene una inercia explicada de un 94,1% por la dimensión 1 y el resto de la inercia la explica la dimensión 2.

Sobre la Contribución de las modalidades a la inercia del factor, el Número de Personas en el Hogar que más contribuyen son 3 y 1 persona en el factor 1 (42,4% y 35% respectivamente) y la que más aporta en el factor 2 es 2 personas en el hogar (70,1%). Luego, la Condición de

desocupado es quien más aporta al factor 1 (89,5 %) y otros ocupados lo hace para el factor 2 (88,7 %).

Con respecto a la Calidad de representación de las modalidades en los ejes, 5 modalidades de Número de Personas en el Hogar están representadas mayormente por el factor 1 (sobre un 96 %) y solo 2 Personas en el Hogar está mejor representada en el factor 2 (98 %). Para la Condición de Actividad, el factor 1 representa mejor a ocupado y desocupado (ambos sobre un 94 %) y el factor 2 representa mejor a otros ocupados (94,8 %).

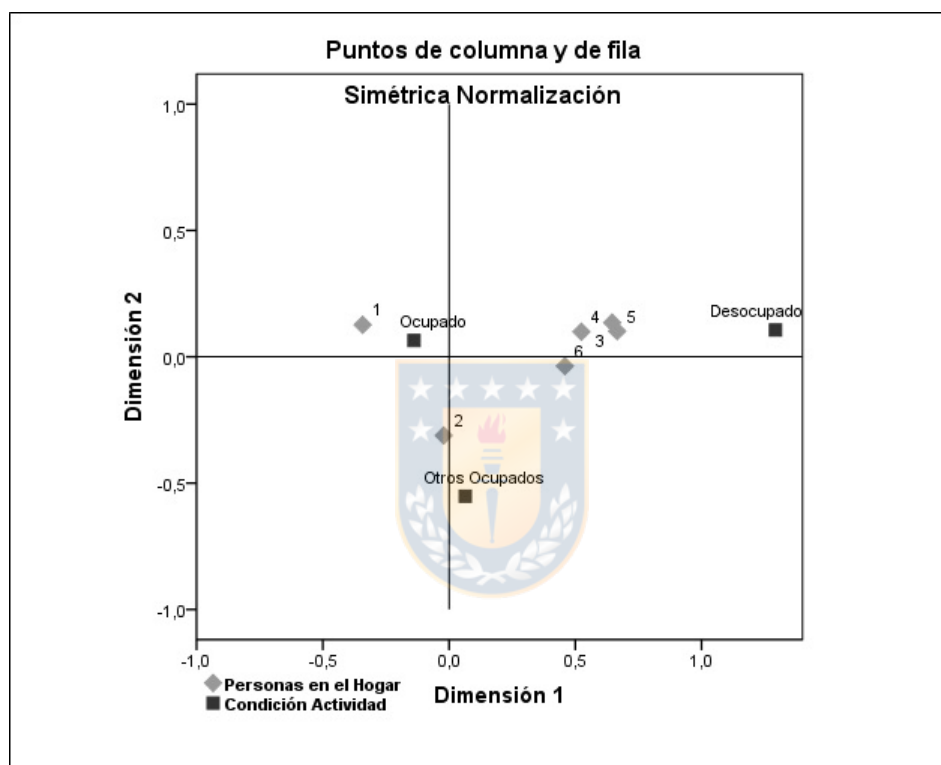


Figura 5.12: ACS de N° de Personas en el Hogar y Condición de Actividad, Noviembre.

De la **Figura 5.12** podemos concluir que la modalidad más similar al perfil medio es la Condición de ocupado y, por el contrario, la modalidad más diferente del perfil medio es la Condición de desocupado. La dimensión 1, al igual que mes anterior, ubica en lado negativo a ocupado y la opone del resto de condiciones y, ubica en mismo lado a una y dos personas en el hogar y las opone del resto de Número de Personas en el Hogar. La dimensión 2 deja la Condición de otros ocupados bajo el eje junto a 2 personas en el hogar y muy levemente bajo el eje a 6 o más personas. Luego, ubicamos en lado derecho y alejado del perfil medio las modalidades que más se asocian, la Condición de desocupado con 3, 4, 5 y 6 o más personas en el hogar. Además, la posición relativa de las modalidades de 3, 4, 5 y 6 o más personas en el hogar son muy cercanas, entonces representan perfiles similares a través de las columnas. Por otra parte, se asocia en lado izquierdo la Condición de ocupado y que la persona viva sola.

Finalmente, se observa un cambio de agosto a noviembre en algunas modalidades del Número de Personas en el Hogar, aunque podrían ser casos particulares o situaciones esporádicas, dado que dichas modalidades mantiene posición en mismo lado que Condición de desocupado, pero una relación más estrecha a la Condición de otros ocupados. Luego, en noviembre hay una marcada tendencia a relacionar la Condición de desocupado y al hecho de que vivan más de 2 personas en el hogar, la relación se mantiene cuando el Número de Personas en el hogar aumenta, cuando son 3, 4, 5 y 6 o más personas en un hogar. Puede ser un síntoma que a fin de año hay mayor desempleo.

Por otra parte, se puede hacer juicio, ya que una persona que viva sola debe estar ocupado y/o con trabajo estable para poder mantener su propia economía y responder a gastos propios de una vivienda. Mientras que, cuando aumenta el Número de Personas en el Hogar se relaciona a ocupados que reciben ingresos de manera no tradicional y también a Condición de desocupado, lo que se puede explicar debido a que no es necesario que todos estén trabajando.

La variable Edad con Condición de Actividad, en mes central de agosto presenta una Calidad de la Representación de la proyección por la dimensión 1 del 79 % de la inercia, y la dimensión 2 explica un 21 % de la inercia total.

Con respecto a la Contribución de las modalidades a la inercia del factor, los rangos de Edad que más contribuyen a la inercia son de 15 a 24 años en el factor 1 (77,3 %) y de 65 o más años en el factor 2 (87,7 %). La Condición de desocupado es quien más contribuye a la inercia del factor 1 (84,9 %) y la Condición de otros ocupados aporta mayormente a la inercia del factor 2 (85 %).

Sobre la Calidad de representación de las modalidades en los ejes, los rangos de Edad de 15 a 64 años están más representados por el factor 1 (sobre 54 %) y el rango de jubilación está mejor representado por el factor 2 (sobre 94 %). La Condición de ocupado y desocupado están más representados por el factor 1 (sobre 85 %), mientras que el factor 2 representa mejor a otros ocupados (92,8 %).

De la **Figura 5.13** podemos observar que las modalidades más similares al perfil medio son la Condición de ocupado y los rangos de 35 a 44 años y de 45 a 54 años. Por el contrario, los puntos más diferentes del perfil medio son desocupado, de 15 a 24 años y de 65 o más años (rango de jubilación). La dimensión 1 muestra los rangos de Edad prácticamente ordenados de izquierda a derecha y, ubica en lado derecho a ocupado que la opone del resto de las condiciones. La dimensión 2 separa los rangos de Edad extremos de los rangos de Edad medios, el rango de 15 a 24 años y el rango de jubilación sobre el eje, también separa por sobre el eje a otros ocupados. Los rangos de Edad entre 35 a 64 años son muy cercanos en el mapa, por ende perfiles de Edad muy similares a través de la Condición de Actividad. En resumen, a la izquierda de la dimensión 1 se relaciona la Condición de desocupado con Edad de 15 a 24 años y no tan próximo pero del mismo lado de 25 a 34 años; también se asocia por sobre la dimensión 2 a otros ocupados con la Edad de 65 o más años y, la Condición de ocupado con los rangos de 35 a 64 años.

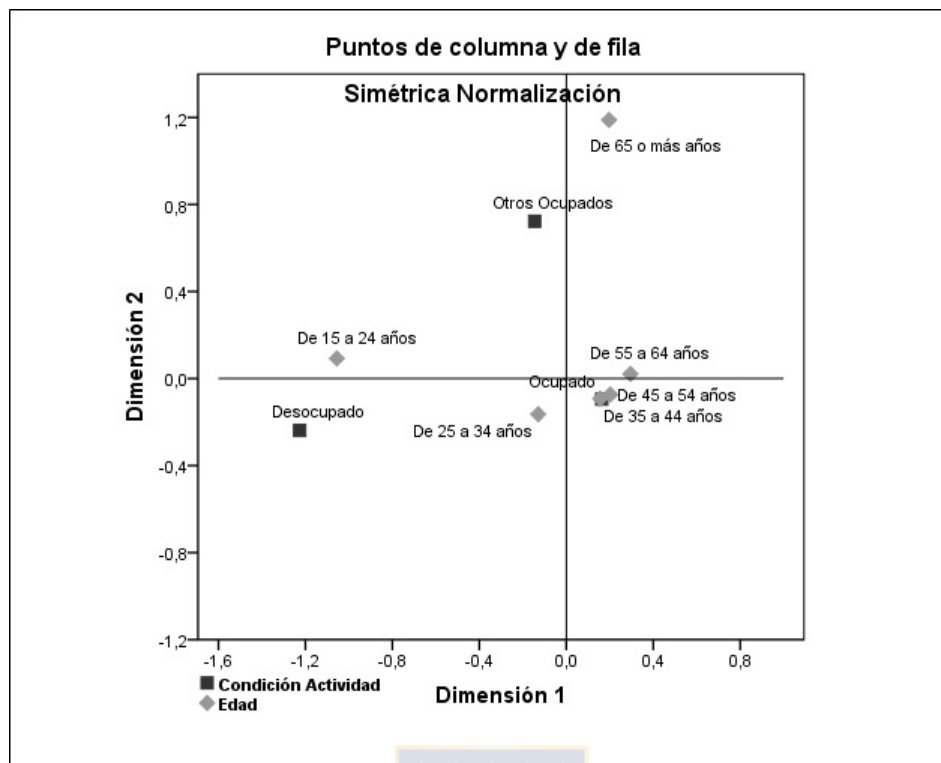


Figura 5.13: ACS de Edad y Condición de Actividad, Agosto.

Las tempranas edades se mantienen relacionadas a Condición de desocupado, dando coherencia a los altos porcentajes de desocupados en Edad temprana obtenido en tablas de contingencia, donde una posible causa principal sería la poca experiencia laboral por representar edad escolar o de formación profesional. Mientras que, los rangos de Edad como potenciales trabajadores activos se mantienen lógicamente asociados a Condición de ocupado. Sorprende que haya una clase de Edad en rangos de jubilación que se mantienen asociados a Condición de otros ocupados, como trabajadores informales o se mantienen teniendo algún tipo de ingresos, sin el común de un contrato laboral.

Para el mes central de noviembre, la Calidad de la Representación de la proyección en la dimensión 1 es de un 90,6 % de la inercia y la dimensión 2 explica el restante 9,4 % de la inercia total.

Sobre la Contribución de las modalidades a la inercia del factor, para la Edad el rango que más contribuye a la inercia es de 15 a 24 años en la dimensión 1 (77,5 %) y el rango de 65 o más años en la dimensión 2 (58,6 %). A su vez, la Condición que más contribuye a la inercia es desocupado en la dimensión 1 (91,4 %) y otros ocupados en la dimensión 2 (88,9 %).

La Calidad de representación de las modalidades en los ejes, para la Edad se tiene que los rangos de 15 a 34 años y de 45 a 64 años están mejor representados por el factor 1 (sobre 73 %) y el factor 2 representa mejor los rangos de 35 a 44 años y de 65 o más años (sobre 56 %). Mientras que, la Condición mejor representada por el factor 1 es ocupado y desocu-

pado (sobre 88%) y otros ocupados está mejor representado en el factor 2 (99%).

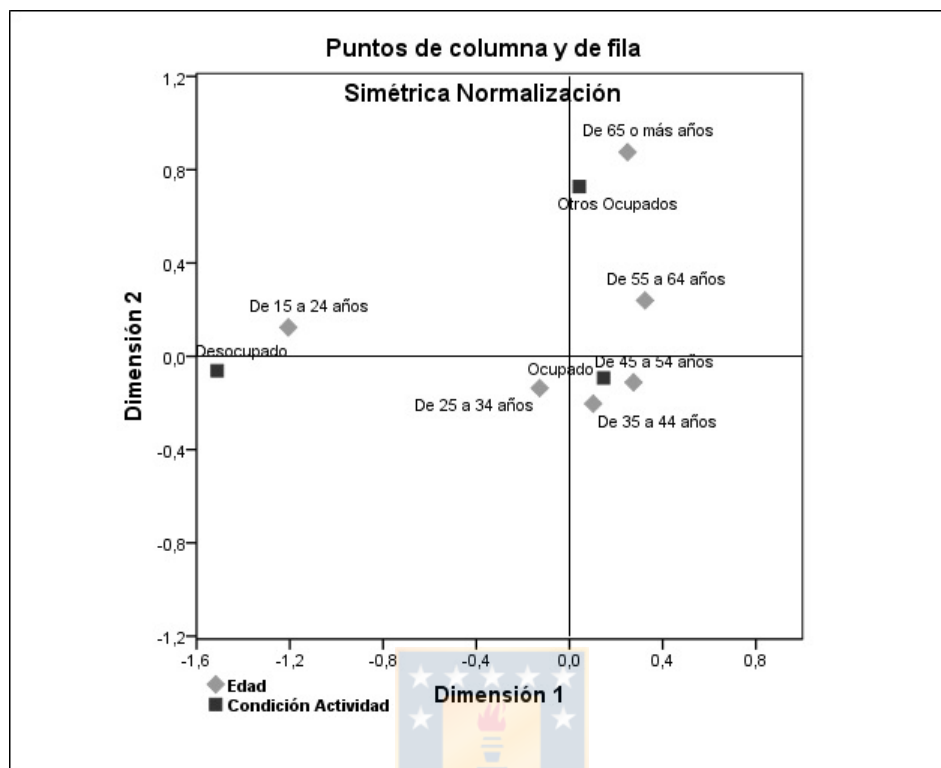


Figura 5.14: ACS de Edad y Condición de Actividad, Noviembre.

De la **Figura 5.14** se observa que, los puntos más similares al perfil medio son ocupado y algunos rangos de Edad entre 25 a 54 años y los puntos más diferentes del perfil medio son desocupado y rangos de 15 a 25 años y la edad de jubilación. La dimensión 1, al igual que mes de agosto, ordena los rangos de edades de izquierda a derecha, con igual lejanía del rango de 65 o más años que el mes anterior. También ubica en lado izquierdo la Condición de desocupado que la opone del resto de las condiciones. La dimensión 2 separa las edades extremas, el rango de 15 a 24 años y los mayores de 55 años sobre el eje y, separa por sobre el eje a otros ocupados. En conclusión, en lado izquierdo de la dimensión 1 se relaciona la Condición de desocupado con el rango de 15 a 24 años; también se asocia la Condición de ocupado con rangos de Edad entre 35 y 54 años; por sobre el eje y a la derecha asoma una relación de otros ocupados con edad de jubilación. Se observa similitud en algunos perfiles de Edad a través de la Condición, por su cercanía, como los rangos de 35 a 44 años y de 45 a 54 años.

Al comparar las conclusiones de agosto a noviembre, se puede resumir que son análogas, hay cambios mínimos que no alteran las principales relaciones. La relación principal y de interés en ambos meses, es la Condición de desocupado con la Edad de 15 a 24 años. Llama la atención una relación que se reitera, de otros ocupados con las Edad de jubilación, por sobre los 65 años.

La siguiente variable en analizar por ACS es **Años de Estudio categorizado con Condición de Actividad**, para el mes central de agosto se tiene que la Calidad de la Representación de la proyección en la dimensión 1 es de un 90 %, mientras que, la dimensión 2 explica el 10 % restante de la inercia total.

Con respecto a la Contribución de las modalidades a la inercia del factor, para Años de Estudio el rango que más contribuye a la inercia es de 0 a 8 años de estudio en la dimensión 1 (71,6 %) y de 19 o más años de estudio en la dimensión 2 (64,4 %). Luego, para la Condición que más aporta a la inercia es otros ocupados en la dimensión 1 (71,2 %) y desocupado en la dimensión 2 (65,4 %).

Sobre la Calidad de representación de las modalidades en los ejes, los rangos de Años de Estudio que mejor representados están por el factor 1 son de 0 a 8 años y de 9 a 12 años de estudio (sobre 91 %), los otros rangos de Años de Estudio están mejor representados en el factor 2 (sobre 61 %). Equivalentemente, para Condición de Actividad resultó que sus 3 modalidades están mejor representadas por el factor 1 (sobre 59 %).

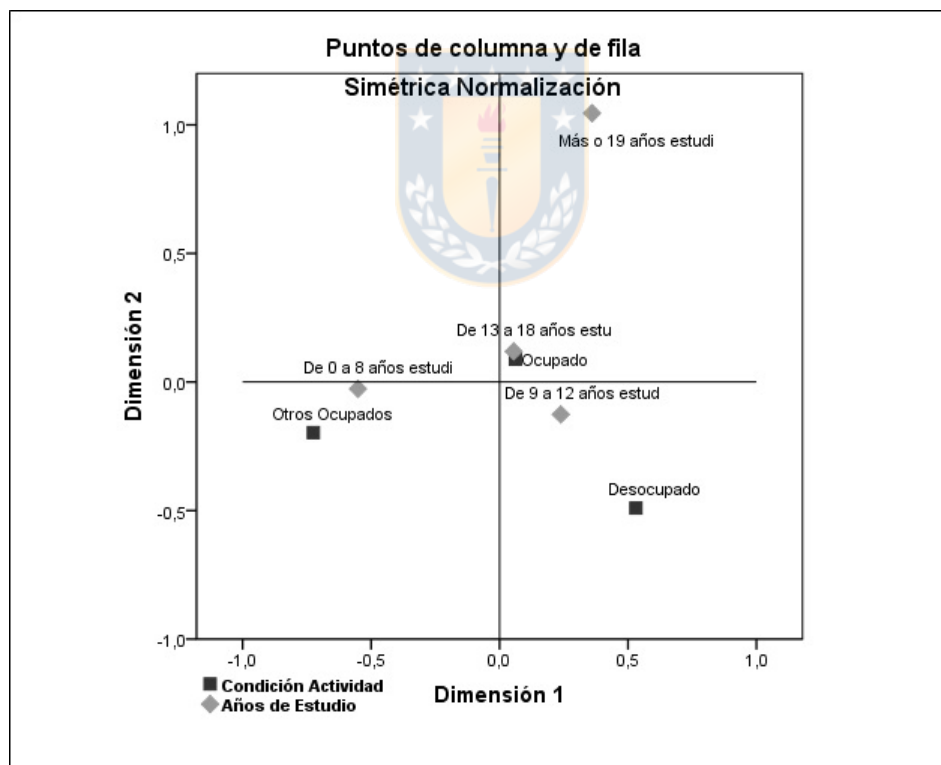


Figura 5.15: ACS de Años de Estudio y Condición de Actividad, Agosto.

De la **Figura 5.15**, se concluye que la modalidad más similar al perfil medio es la Condición de ocupado y la clase de 13 a 18 años de estudio. Por el contrario, la modalidad más diferente del perfil medio es la clase de 19 o más años de estudio. La dimensión 1 ubica en lado izquierdo el rango de 0 a 8 años de estudio y la Condición de otros ocupados, el resto

de modalidades se ubican en lado derecho. La dimensión 2 separa los Años de Estudio en rango de estudio superior de los Años de Estudio de educación básica y media, además deja la Condición de ocupado sobre el eje, con los rangos de estudio superior. Por último, en lado izquierdo del factor 1, se relaciona la Condición de otros ocupados con el rango de 0 a 8 años de estudio, en lado derecho y bajo el eje horizontal se relaciona la Condición de desocupado con el rango de 9 a 12 años de estudio, además, la Condición de ocupado con el rango de 13 a 18 años de estudio.

El mes central de noviembre, para la variable Años de Estudio, tiene una Calidad de la Representación de la proyección en la dimensión 1 del 58,2%, y la dimensión 2 explica el 41,8% restante de la inercia total.

Con respecto a la Contribución de las modalidades a la inercia del factor, los rangos de Años de Estudio que más contribuyen a la inercia son de 0 a 8 años de estudio en el factor 1 (59%) y de 13 a 18 años de estudio en el factor 2 (55,3%). Por otra parte, la Condición que más aporta a la inercia del factor 1 es desocupado (87,6%) y al factor 2 es la Condición de otros ocupados (80,1%).

Además, la Calidad de la representación de las modalidades en los ejes para Años de Estudio, resultó con casi todos los rangos mejor representados por el factor 1 (sobre 55%), solamente el rango de 13 a 18 años de estudio está mejor representada por el factor 2 (90,6%). De forma equivalente, la Condición de desocupado está mejor representada por el factor 1 (96,6%) y las otras condiciones están mejor representadas por el factor 2 (sobre 76%).

De la **Figura 5.16** se concluye que, las modalidades más similares al perfil medio son la Condición de ocupado y el rango de 9 a 12 años de estudio, por el contrario, las modalidades más diferentes del perfil medio son la Condición de desocupado y el rango de 19 o más años de estudio. La dimensión 1 ubica en lado negativo la Condición de desocupado y al resto de las condiciones en lado positivo y, separa en mismo lado negativo los rangos de 9 a 12 años y de 13 a 18 años de estudio del resto de Años de Estudio, que son los rangos extremos. La dimensión 2 tiene una representación semejante al mes anterior, separa los Años de Estudio de educación superior sobre el eje de los Años de Estudio de educación básica y media bajo el eje, también ubica la Condición de ocupado sobre el eje y las otras condiciones bajo el eje. En conclusión, se relaciona en lado izquierdo y bajo el eje la Condición de desocupado y el rango de 9 a 12 años de estudio, en lado derecho y bajo el eje a otros ocupados con el rango de 0 a 8 años de estudio, finalmente, la Condición de ocupado se relaciona con el rango de 13 a 18 años de estudio.

La variable Años de Estudio también presenta un comportamiento idéntico en mes de agosto y noviembre, las relaciones se mantienen y como es de suponerse la Condición de ocupado se relaciona mayormente con Años de Estudio de educación superior. Por otra parte, una relación muy marcada es para la Condición de desocupado con los rangos de 9 a 12 años de estudio, en ambos meses.

Esto implica que, las personas que completaron la enseñanza media o no la completaron, son

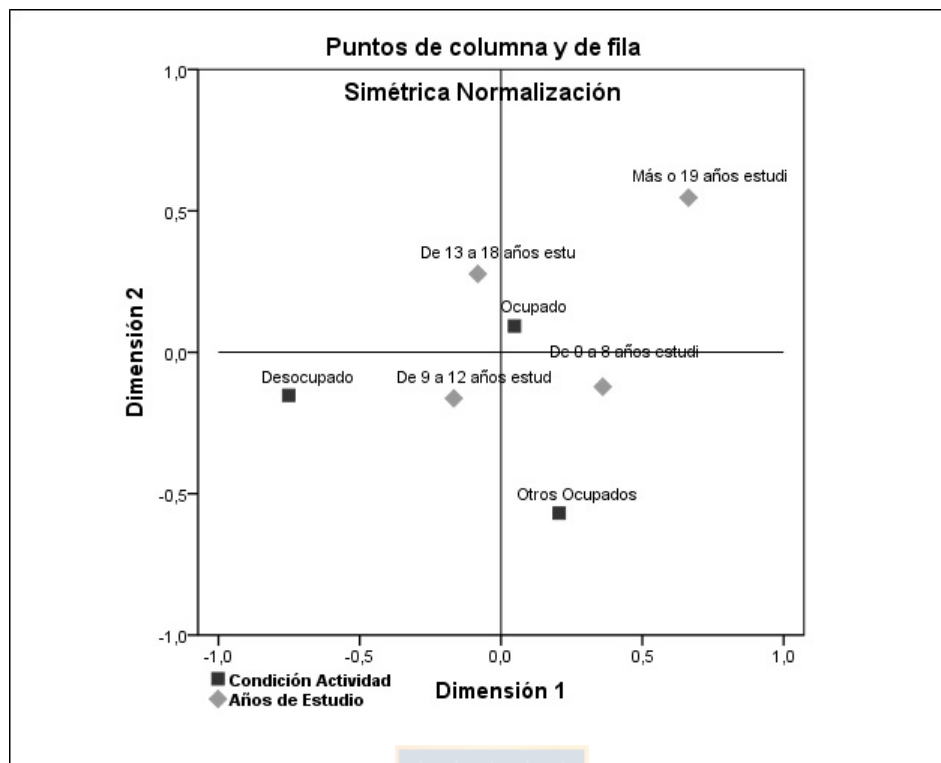


Figura 5.16: ACS de Años de Estudio y Condición de Actividad, Noviembre.

más susceptibles a no tener empleo constante y dada su condición de no tener especialización en algún área de trabajo, se mantienen en el tiempo mayormente desempleados.

Finalmente, la comparación realizada de los ACS de mes central de Agosto con los ACS de mes central de Noviembre en la Región del Biobío, terminó por confirmar las mismas relaciones para la Condición de desocupado. No hubo cambios drásticos o modificaciones evidentes en los mapas analizados, salvo algunas modalidades que se alejaban o se aproximaban a la Condición de interés. Sin embargo, la principal relación encontrada en la Condición de desocupado junto a cada variable en estudio se mantuvo en ambos meses, evidenciando o confirmando un idéntico comportamiento de los registros de individuos. Se puede concluir poca variación en las estaciones en estudio para la Región del Biobío.

Los ACS realizados, además actuaron como confirmación de los mayores porcentajes de cada modalidad de las variables en estudio, dada la Condición de desocupado. Salvo unas excepciones en variables como Grupo de Ocupación, Rama de Actividad y Número de Personas en el Hogar, las que no mostraron en estricto rigor como una relación principal la Condición de desocupado con el mayor porcentaje de las modalidades de esas variables.

Ya se observa una estructura a modo bivariado de la Condición de desocupado con individuos de Provincia de Concepción y de Estrato de ciudades, el Sexo presentó un comportamiento similar junto a la Condición de ocupado relacionándolos con la categoría hombre. También se

encontró relación de la Condición de desocupado con encuestados soltero(a), con individuos no proveedor, con individuos principalmente de Grupo de Ocupación 3 (ocupaciones elementales o trabajadores no calificados como asistentes, peones, ayudantes, vendedores ambulantes, recolectores, otros no identificados), aunque el Grupo 2 de Ocupación también estaba próximo a desocupado. Otras relaciones de forma bivariada encontradas con la Condición de desocupado fue con encuestados que tuvieron empleos en Rama de Actividad económica secundaria (explotación de minas y canteras, industrias manufactureras, suministro de electricidad, gas y agua, y construcción), con individuos que tuvieron empleos dependiente, con encuestados en hogares numerosos de 3 a 6 o más personas, individuos con Edad de 15 a 24 años principalmente (también, mayormente próximo a desocupado que las otras condiciones, el rango de 25 a 34 años) y, una relación de desocupado con individuos de 9 a 12 Años de Estudio.

5.4. Análisis regional mediante ACM.

En el siguiente análisis se presentan **los resultados de Análisis de Correspondencia Múltiple (ACM)**, para el conjunto completo de las 11 variables independientes en estudio más la variable de interés Condición de Actividad, para cada mes central. El objetivo es comparar los ACM de los meses centrales y, encontrar una estructura de asociación para la condición de desocupado en los encuestados de la Región del Biobío. Se realizó el ACM para cada mes central, se observó y analizó: las bajas inercias de ambos análisis como dice la teoría, por lo que se procedió a realizar la Corrección de Benzecri para realizar una interpretación de la calidad de representación de la proyección; las medidas de discriminación; y los mapas de categorías conjuntas para 3 dimensiones en ambos meses, para la interpretación se tomaron combinaciones de dimensiones de a pares, para ver solo mapas bidimensionales.

La cantidad de variables para el ACM son 12, con un total de 41 modalidades, por lo tanto, bajo la fórmula de inercia total definida (capítulo 2 sección 2.3.2) se obtiene una inercia total de $(41 - 12)/12 = 2,4167$. Para la corrección de Benzecri, en las inercias principales, se utilizó las dimensiones cuya inercia principal eran superior a $1/12 = 0,0833$.

Con respecto al mes central de Agosto, se analizó y se realizó la Corrección de Benzecri para las inercias principales, en la **Tabla 5.10** se observa las inercias principales retenidas (correspondientes a las de una Tabla Disyuntiva Completa), las inercias ajustadas o corregidas, porcentaje de inercia explicada y porcentaje de inercia explicada acumulada.

Se puede corroborar la baja inercia explicada por cada dimensión en las inercias principales originales, la primera no supera una inercia principal de 0,25 del total de inercia 2,4167.

Con respecto a la Corrección de Benzecri, para explicar la representación de la proyección en ACM, se tiene que la primera dimensión explica un 62,53 % de la inercia ajustada, le sigue la segunda dimensión con un 19,98 % de la inercia corregida y la dimensión 3 que explica un

8,14%. Las 3 dimensiones acumulan una inercia ajustada o corregida explicada del 90,65%. Por lo tanto, se observó la gráfica de dimensión 1 vs dimensión 2, una segunda gráfica de la dimensión 1 vs dimensión 3 y, como último gráfico de categorías conjuntas la dimensión 2 vs la dimensión 3.

Dimensión	Inercia TDC	Corrección Benzecri	% Inercia explicada	% Inercia acumulada
1	0,24964	0,0329155	62,53153	62,53153
2	0,17734	0,0105165	19,97883	82,51037
3	0,14333	0,0042840	8,13849	90,64885
4	0,12971	0,0025593	4,86199	95,51084
5	0,11210	0,0009847	1,87077	97,38160
6	0,10887	0,0007764	1,47492	98,85653
7	0,10148	0,0003918	0,74430	99,60082
8	0,09409	0,0001378	0,26181	99,86264
9	0,08930	0,0000423	0,08036	99,94299
10	0,08734	0,0000191	0,03624	99,97924
11	0,08601	0,0000085	0,01621	99,99545
12	0,08475	0,0000024	0,00455	100,000
Suma	1,463957	0,0526382	100,000	

Tabla 5.10: Corrección de Benzecri para inercias de ACM en Agosto.

VARIABLES	Dimensión 1	Dimensión 2	Dimensión 3
Provincias	0,0688	0,3949	0,0701
Tipo Estrato	0,1040	0,4877	0,0560
Personas en el Hogar	0,5726	0,1736	0,2833
Edad	0,4781	0,1442	0,0961
Sexo	0,1172	0,0119	0,4342
Años de Estudio	0,3117	0,1814	0,0949
Estado Conyugal	0,3798	0,1323	0,0687
Proveedor	0,5132	0,1254	0,0393
Grupo de Ocupación	0,1342	0,1776	0,0863
Rama de Actividad	0,1857	0,2370	0,3223
CISE	0,0864	0,0229	0,0582
Condición Actividad	0,0441	0,0393	0,1108
Promedio (Inercia)	0,2496	0,1773	0,1433

Tabla 5.11: Medidas de discriminación de ACM en Agosto.

Las Medidas de Discriminación se observan en la **Tabla 5.11**. Entonces, la variable Número de Personas en el Hogar, Edad, Años de Estudio, Estado Conyugal y Proveedor están muy relacionadas con el factor 1. Además, la variable Provincia y Tipo Estrato están más relacionadas al factor 2, y, la variable Sexo y Rama de Actividad tienen mayor relación al factor 3. Las variables CISE, Grupo de Ocupación y Condición de Actividad, tienen una

baja relación con los factores, pero están mayormente representadas en los factores 1, 2 y 3 respectivamente.

Para la dimensión 1 (D1) vs la dimensión 2 (D2), se observa primero la **Figura 5.17** de representación de las variables en ACM. Se encuentran graficadas las medidas de discriminación para la dimensión 1 y 2. El gráfico muestra que las variables Número de Personas en el Hogar, Proveedor, Edad, Estado Conyugal y Años de Estudio están más relacionadas al factor 1. Mientras que, Provincia y Tipo Estrato lo están más al factor 2. Además, las variables Grupo de Ocupación y Rama de Actividad tienen similar relación con ambos factores. Por último, se observa una baja relación de las variables Sexo, CISE y Condición de Actividad con los factores.

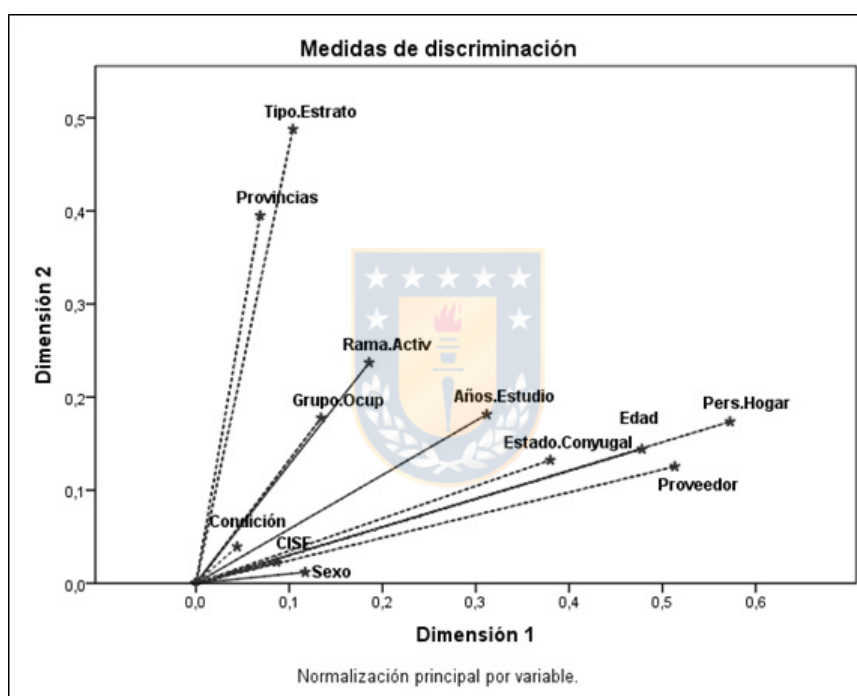


Figura 5.17: Representación de las variables en ACM, D1 vs D2 Agosto.

Se presenta el diagrama conjunto de todas las modalidades de las 12 variables en la **Figura 5.18**, los resultados del ACM. Cada símbolo representa una modalidad de cada variable y la etiqueta de cada modalidad tiene asociada un número de 1 hasta la cantidad de categorías de cada variable, que corresponde a la categorización realizada en la sección de “variables utilizadas”. Por último, la leyenda presenta el formato de cada variable. Se observa que hay modalidades alejadas del origen, lo que representa su poca frecuencia en los registros, por ejemplo: para Años de Estudio la modalidad 4, se ubica por bajo el eje 2, de 19 o más años de estudio que representa un 1,9% de los encuestados (100 encuestados, Tabla 5.5) o, para Rama de Actividad la modalidad 1, ubicada por sobre el eje y alejada hacia la izquierda, actividad primaria que representa a un 7,1% de los registros (369 encuestados, Tabla 5.3). Por el contrario, hay modalidades de mucha frecuencia, son las ubicadas más cercas del ori-

gen, por ejemplo: en Condición de Actividad la modalidad 1, que corresponde a encuestados ocupados; para Edad la modalidad 3, son encuestados de 35 a 44 años; en la variable Años de Estudio la modalidad 2, corresponde a encuestados con 9 a 12 años de estudio; y la variable Clasificación Internacional de Categoría del Empleo la modalidad 2 que son personas con trabajos dependientes.

Se observa proximidad de modalidades de una misma variable, como el Número de Personas en el Hogar que están agrupadas las clases 3, 4, 5 y 6 en lado positivo del factor 1 y por sobre el eje, esto quiere decir que el grupo de encuestados que dicen vivir 3, 4, 5 y 6 o más personas en el hogar son semejantes, entonces tienen casi el mismo comportamiento en las otras variables. De igual forma, la variable Estado Conyugal presenta cercanía en 2 de sus modalidades en la parte inferior izquierda, la categoría 1: casado(a) y 4: viudo(a), separado(a)/anulado(a) o divorciado(a), lo que implica semejanza entre los individuos casados(as) y los viudos(as) o divorciados(as), ya que los grupos que eligieron estas modalidades tiene similar comportamiento en las otras variables. Otro caso equivalente sucede en la Edad, la cual tiene 2 modalidades cercanas, 4 y 5, que corresponden a rangos de edad de 45 a 54 años y de 55 a 64 años, los grupos de individuos en este rango etario son semejantes dado su similar comportamiento en las otras variables. Por último, Grupo de Ocupación en sus modalidades 2 y 3 también se encuentran muy cercanos, sin embargo, su relación con los factores es algo baja.

Sobre las variables relacionadas al factor 1, se observa que Edad está estructurada a lo largo del eje de derecha a izquierda (de 1 a 6), las modalidades 1 y 2 sobre el eje y a la derecha del mapa, el resto de modalidades se ubica a la izquierda del mapa comenzando con modalidad 3 cercana al origen y terminando con la clase 6 (de 65 o más años) siendo la más alejada del centro. La variable Número de Personas en el Hogar tiene un comportamiento similar, pero del cuadrante inferior izquierdo al cuadrante superior derecho del mapa. Además, estas 2 variables son similares, ya que se encuentran muy cercanas a lo largo del factor 1. Estado Conyugal separa sus modalidades semejantes en la parte inferior izquierda (donde se encuentran próximas a rangos etarios sobre 45 años), a la parte superior donde se encuentra 2: conviviente y 3: soltero(a), más próximos a rangos etarios jóvenes. Además, Proveedor de 2 modalidades se encuentra opuesta, con respecto al origen, proveedor bajo el eje a la izquierda y no proveedor sobre el eje a la derecha. La variable Años de Estudio se estructura de izquierda a derecha, con tendencia hacia la parte inferior mientras más años de estudio tiene el individuo, la clase 1 es de 0 a 8 años de estudio, la clase 2 cercana al origen con 9 a 12 años de estudio, la clase 3 con 13 a 18 años de estudio a la derecha del factor 1 y, su última modalidad (4: de 19 o más años de estudio) es la más alejada y verticalmente bajo el eje.

Con respecto a las variables relacionadas al factor 2, principalmente Provincia y Tipo de Estrato, se ubican en la extensión del eje vertical, bajo el eje y cercano al origen se encuentran próximas la provincia de Concepción y el estrato ciudades, muy alejadas del origen en la parte superior izquierda se encuentra provincia de Arauco y área rural, y restos de áreas urbanas se encuentra entre la provincia de Arauco y Bío Bío. Del resto de variables, se comenta que poseen una baja relación con cada factor pero Grupo de Ocupación y Rama de Actividad se estructuran de forma similar del cuadrante superior izquierdo al cuadrante

inferior derecho; las variables que poseen 2 modalidades como Sexo y CISE tienen aún más baja relación con los 2 factores y se ubican de forma opuesta con respecto al origen; y la variable Condición de Actividad también se encuentra muy poco relacionada a los 2 factores.

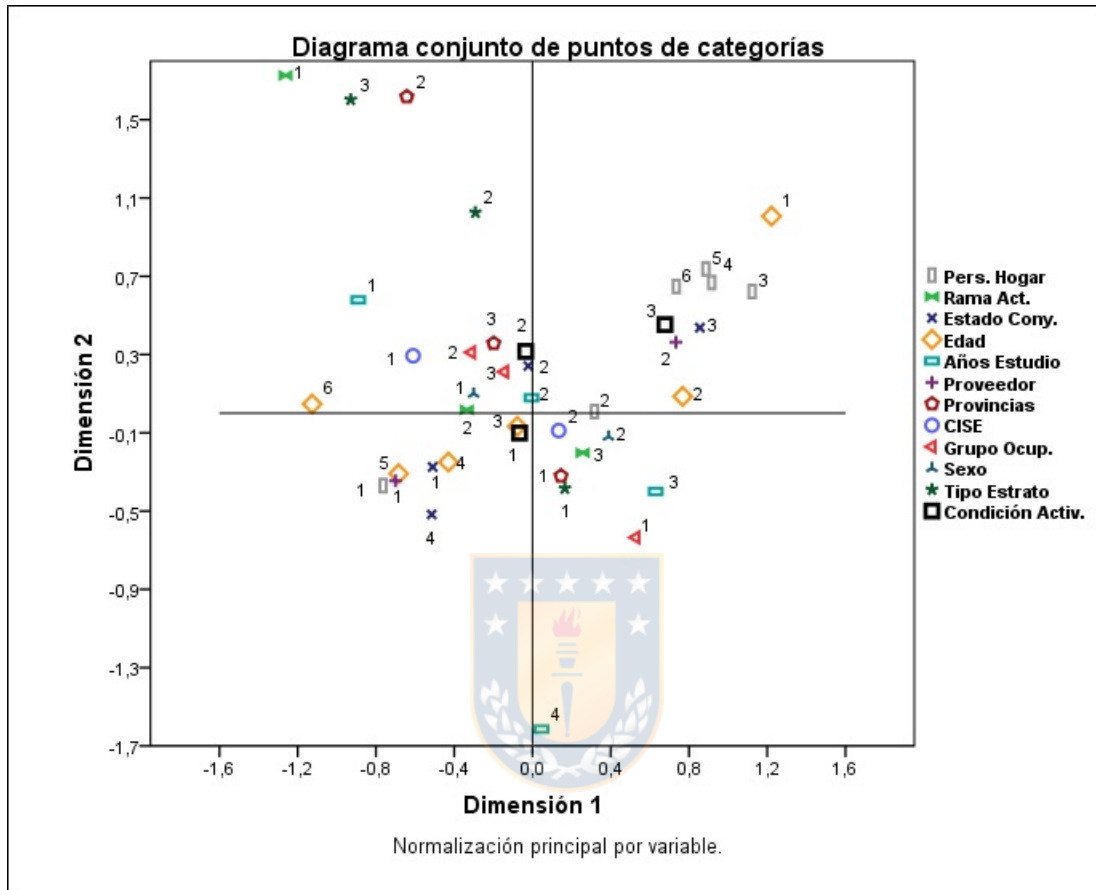


Figura 5.18: ACM de todas las variables, D1 vs D2 Agosto.

Sobre la proximidad de modalidades de distintas variables, en particular referente a las variables más relacionadas al factor 1, se observa a izquierda de la dimensión 1 rangos etarios mayor a 35 años, individuos casado(a) y viudo(a) o divorciado(a), el grupo de proveedor que pueden vivir solos en el hogar y hasta 12 años de estudio, por lo que existe una relación entre ellos. Además, las condiciones de Ocupados y otros ocupados, a pesar de su poca relación a las dimensiones, se encuentran ubicados en el lado izquierdo de esta. Se destaca en la parte superior de la dimensión 1, la actividad primaria junto a área rural y Provincia de Arauco, próximo a ellas se encuentra resto de áreas urbanas (RAU) le sigue sin estudio o hasta 8 años de estudio, lo que indica una relación entre ellas. Algo esperada en las actividades de agricultura, ganadería, caza y silvicultura o pesca en estratos que no sean grandes ciudades, además relacionados a individuos con hasta 8 años de estudio, en este mismo cuadrante se ubica cerca del eje horizontal el grupo etario de 65 o más años (tercera edad).

Por otra parte, se observa a la derecha y bajo el eje los individuos que poseen 13 o más

años de estudio, representando individuos con estudios superiores y de post grado, el Grupo de Ocupación 1 (directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio), los individuos de la Provincia de Concepción y de estrato Ciudades, que representan una relación entre ellas, algo lógica dada la ubicación geográfica de mayores Universidades y centros de estudios superiores, junto a dichas ocupaciones y próximas también a actividades terciarias. Luego, por sobre el eje en el mismo lado derecho se observan próximos rangos etarios jóvenes hasta 34 años, que pueden vivir de 3 a 6 o más personas en el hogar, individuos solteros(as) y no proveedor, junto a la Condición de desocupado, lo que implica una asociación entre ellas dado que han sido elegidas por los mismos individuos o presentan individuos parecidos con respecto a las otras variables. Aunque la Condición de Actividad Desocupado no se encuentra muy relacionada a estas dimensiones, se ubica perfectamente en este grupo, y se observa que el factor 1 (que más explica de la inercia corregida) opone a las condiciones de ocupado y otros ocupados a la izquierda de la dimensión 1, de los individuos desocupados en la derecha del mapa.

Para la dimensión 1 (D1) vs la dimensión 3 (D3), se observa la **Figura 5.19** con las medidas de discriminación del ACM. Claramente se observan las mismas variables relacionadas al factor 1. Mientras que, la variable Sexo y Rama de Actividad asumen mayor relación con el factor 3. Luego, las variables Provincia, Tipo de Estrato y grupo de Ocupación pierden relación con estas dimensiones; las variables CISE y Condición de Actividad mantienen una baja relación con los factores, pero en esta combinación de dimensiones es la mejor relación que logran.

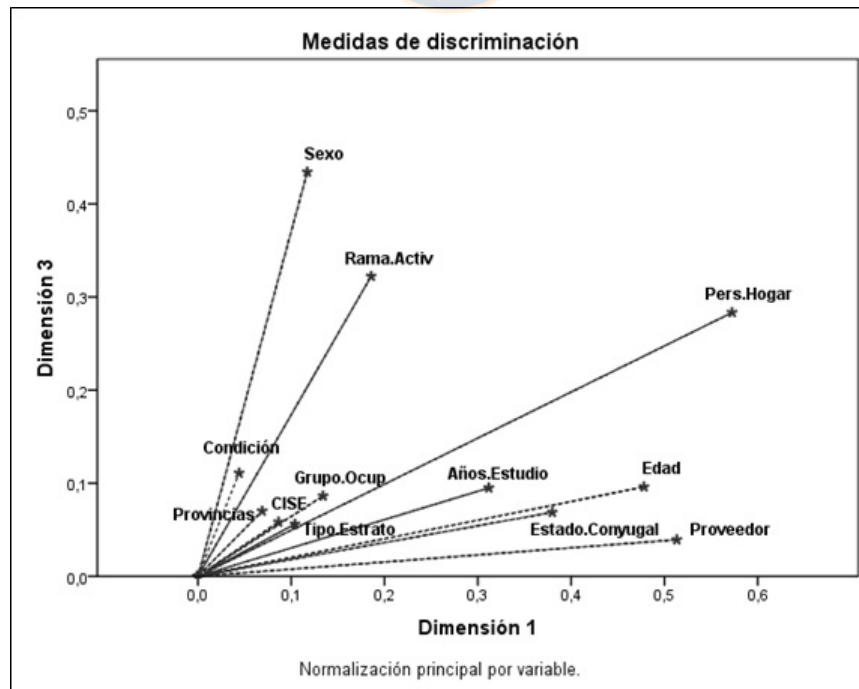


Figura 5.19: Representación de las variables en ACM, D1 vs D3 Agosto.

En el diagrama conjunto de las modalidades de la dimensión 1 vs la dimensión 3, la **Figura 5.20**, mantiene las mismas etiquetas y el mismo formato anterior (Figura 5.18). Se puede observar que las modalidades alejadas del origen no están muy claras, por ejemplo: la variable Edad en su modalidad 1, que aparece a la derecha y bajo el eje, pero próxima a otras modalidades; para Rama de Actividad se repite actividades primarias a la izquierda del factor 1. Por el contrario, se observan modalidades muy cerca del origen (de mucha frecuencia) como Condición de ocupado, Edad de 35 a 44 años y CISE dependiente, las mismas que análisis anterior, pero se suman las modalidades de Estado Conyugal conviviente, Provincia de Concepción y Estrato ciudades (éstas 2 últimas no muy bien relacionadas a las dimensiones).

Con respecto a las cercanías de modalidades de una misma variable, coincide el Número de Personas en el Hogar con sus clases 3, 4, 5 y 6, que son individuos que dicen vivir 3, 4, 5 y 6 o más personas en el hogar, pero esta vez se ubican bajo el eje horizontal y hacia la derecha; la variable Edad mantiene una leve cercanía entre las mismas modalidades con respecto al análisis anterior, de 45 a 54 años y de 55 a 64 años. Estas cercanías hablan de semejanzas entre individuos que tienen esas características en Edad y Número de Personas en el Hogar, ya que mantienen un comportamiento similar en el resto de las variables.

De las variables más relacionadas con el factor 1, se mantiene la estructura ordenada de la variable Edad de derecha a izquierda, comenzando bajo el factor 3 y terminando con la modalidad 6 sobre el factor 3. La variable Número de Personas en el Hogar se presenta en forma de parábola vertical, comenzando de 1 persona bajo el eje y a la izquierda del factor 1, terminando con el cúmulo de clases de 3 a 6 o más personas en el hogar bajo el eje y a la derecha del factor 1, la modalidad 2 se encuentra en la parte superior del factor 3. La variable Años de Estudio se estructura en una forma de elipse horizontal, ordenando sus primeras modalidades (de 0 a 18 años de estudio) de izquierda a derecha y, la modalidad 4: de 19 o más años de estudio por sobre el eje horizontal pegada al eje vertical. Estado Conyugal ubica a soltero(a) en lado positivo del factor 1 bajo el eje y la opone del resto de modalidades, y Proveedor se ubica en forma simétrica, pero cambiando de posición relativa dada su baja representación en la dimensión 3.

Por otra parte, las variables más relacionadas con el factor 3, como Sexo se ubica de manera opuesta sobre el origen, dejando la modalidad 1: hombre en la parte negativa del factor 3 lado izquierdo y modalidad 2: mujer en la parte positiva del factor 3 y a la derecha; además, Rama de Actividad ubica la actividad secundaria bajo el eje, muy cerca de hombres, y las modalidades 1 y 3, por sobre el eje horizontal, actividades terciarias próxima a mujeres. Condición de Actividad tiene su mejor representación en este factor, aunque sigue siendo baja, ubica a desocupado en cuadrante inferior derecho y sus complementos en cuadrante superior izquierdo, con la modalidad 1 de ocupado muy cerca del origen y sobre el eje vertical.

Finalmente, se analiza el mapa observando cercanía de modalidades de distintas variables, se mantiene en lado izquierdo de la dimensión 1 los rangos etarios mayor a 35 años, individuos casado(a) y viudo(a) o divorciado(a) que son proveedor, que pueden vivir solo en el hogar con hasta 12 años de estudio y la Condición de ocupado y otros ocupados; pero en cuadrante negativo del factor 3 se distinguen individuos de Sexo hombre con actividad secundaria y

que pueden vivir solos en el hogar y son proveedor (una estructura algo lógica), sus años de estudio son de 9 a 12 años, por lo que se encuentra esta nueva relación en este grupo. Además, sobre el eje se evidencia mejor la relación de actividad primaria junto a Edad de jubilación y el rango de 55 a 64 años, individuos sin estudio o hasta 8 años de estudio y una Categoría del Empleo como independiente. A pesar del poco aporte a estas dimensiones de Provincia y Tipo Estrato, se mantienen junto actividad primaria y este grupo el área rural y Arauco, como en análisis anterior.

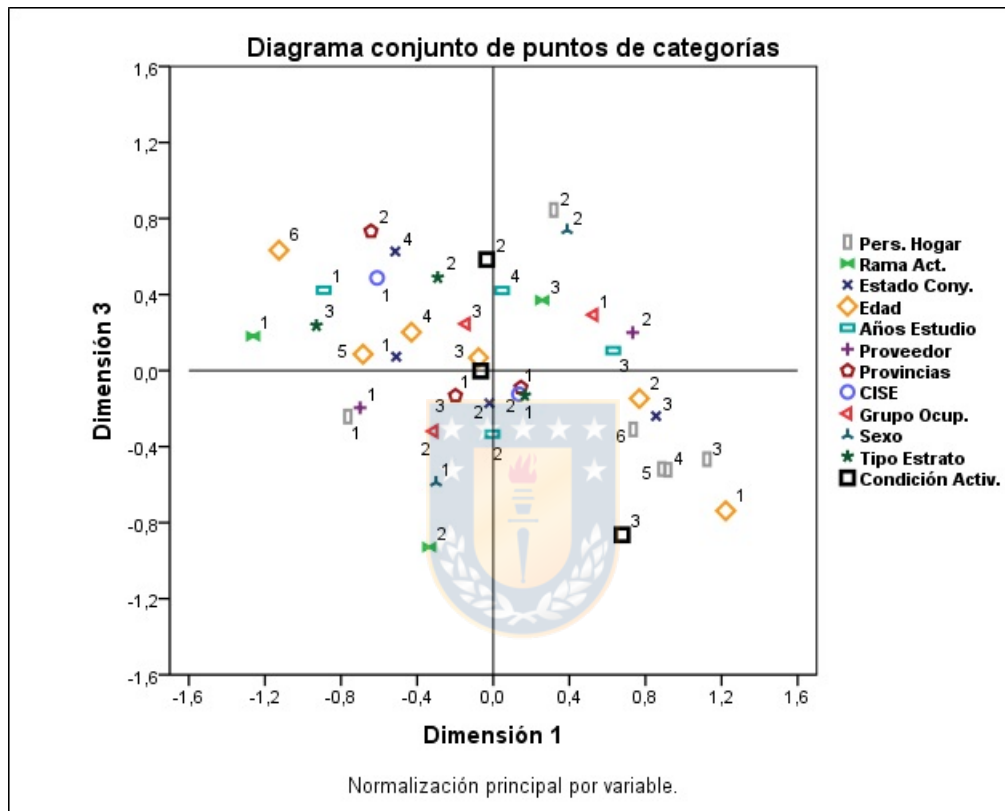


Figura 5.20: ACM de todas las variables, D1 vs D3 Agosto.

Por el contrario, en lado positivo de la dimensión 1 y sobre el eje se observa individuos de Sexo mujer próximo a un hogar con 2 personas y a actividad terciaria, representando una nueva relación. Además, se encuentran cercanos y se mantienen individuos con 13 o más años de estudio y Grupo de Ocupación 1 (a pesar que disminuyen su aporte en factor 3) relacionando estos grupos. Luego, bajo el eje se mantienen rangos etarios jóvenes hasta 34 años, hogares de 3 a 6 o más personas, individuos solteros(as) junto a la Condición de desocupado, lo que indica que se mantiene esta relación, sin incluir a no proveedor que disminuye su representación en el factor 3, y recordando que el aporte de Condición de Actividad es el mejor en estas dimensiones. La CISE igual tiene su mejor representación, el Empleo dependiente se ubica en el mismo cuadrante pero cercana al origen (por su alta frecuencia), junto a ella, la Provincia de Concepción y el Estrato ciudades ambas en su más baja representación.

Para la dimensión 2 (D2) vs la dimensión 3 (D3), se observa la **Figura 5.21** con las medidas de discriminación del ACM. En el factor 2 se observa la relación con las variables Provincia, Tipo de Estrato y Grupo de Ocupación en su mayor aporte. A su vez, las variables Sexo, Rama de Actividad y Número de Personas en el Hogar evidencian su relación con el factor 3. En este análisis, pierde relación con los factores las variables Edad, Años de Estudio (quedando en igual aporte que Grupo de Ocupación), Estado Conyugal y Proveedor; cae aún más su relación con los factores la variable CISE. La variable Condición de Actividad tiene su mayor aporte en el factor 3, aunque se recuerda que es baja.

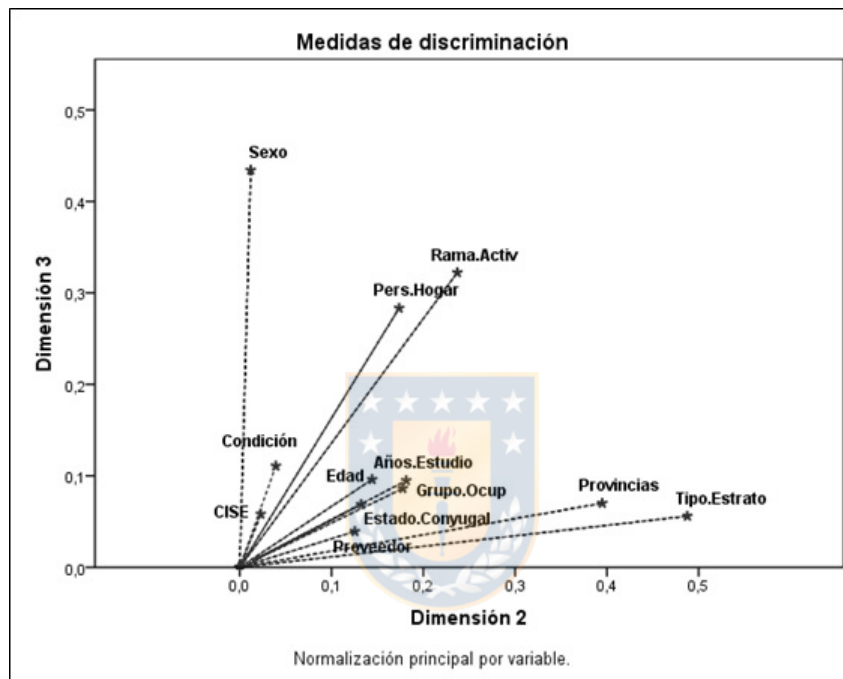


Figura 5.21: Representación de las variables en ACM, D2 vs D3 Agosto.

El diagrama conjunto de modalidades para estas nuevas dimensiones, se observa mediante la **Figura 5.22**, manteniendo las mismas etiquetas y formato.

Las modalidades alejadas del origen se repiten con las del primer análisis, dada la representación de la dimensión 2, de 19 o más años de estudio y actividad primaria, y muy cercana se encuentra área rural, se debe principalmente a su poca frecuencia. Por el contrario, las modalidades cerca del origen se repiten, como lo son la Condición de ocupado, el rango de Edad de 35 a 44 años, la CISE dependiente, influidas principalmente por su alta frecuencia.

Sobre las proximidades de modalidades de una misma variable sucede lo mismo, se reitera el Número de Personas en el Hogar que van de 3 a 6 o más personas, ubicadas en el cuadrante inferior derecho y, la variable Edad con sus rangos de 45 a 54 años y de 55 a 64 años, que en este análisis se ubican cerca la modalidad 3: de 35 a 44 años. Mantienen esta proximidad a pesar de haber disminuido su aporte y representatividad en las dimensiones correspondientes. Una nueva cercanía se observa entre conviviente y soltero(a), de la variable Estado

Conyugal, pero en este análisis la variable bajó su relación con las dimensiones.

En cuanto a las variables más relacionadas al factor 2, Provincia y Tipo de Estrato se encuentran estructuradas cerca del origen en lado negativo hasta llegar al lado derecho alejado del origen. El Grupo de Ocupación, en su mayor relación con un factor, ubica a grupo 1 (Directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleados de oficina) en lado negativo y los grupos 2 (bajo el eje) y 3 (sobre el eje) en lado derecho. Años de Estudio por su parte, se estructura de derecha con 0 a 8 años de estudio a izquierda, con 19 o más años de estudio, se ubica bajo el eje solo de 9 a 12 años de estudio. De forma análoga, las variables más relacionadas al factor 3, Sexo se opone de forma vertical (dado sus 2 modalidades), Número de Personas en el Hogar se presenta en forma de parábola vertical idéntico que análisis anterior, Rama de Actividad ubica a actividad secundaria bajo el eje y la opone del resto de actividades igual que análisis anterior. Condición de Actividad en su mejor representación, pero con bajo aporte a las dimensiones, ubica a desocupado bajo el eje horizontal y por sobre el eje literalmente a ocupado (en lado izquierdo) y a otros ocupados más alejado (en lado derecho).

En conclusión, observando el mapa en su conjunto, en lado izquierdo de la dimensión 2 se ubica rangos etarios de 35 a 64 años (edad de jubilación se ubica levemente en la derecha), individuos casado(a) y viudo(a) o divorciado(a) que son proveedor y viven solos, manteniendo una asociación a pesar que han disminuido su representación en estas dimensiones. Bajo el eje se distingue una nueva asociación en individuos de Concepción, obviamente de estrato ciudades y que viven solo, siendo más débil que las anteriores. Sobre el eje, se mantiene una asociación entre actividad terciaria, Sexo mujer, grupo de ocupación 1 y de 13 o más años de estudio (educación superior y posgrados). Ha mantenido una asociación el rango de Edad de 35 a 44 años junto a la Condición de ocupado, en los 3 análisis.

En lado derecho de la dimensión 2, por sobre el eje, se destaca que mantiene una proximidad la Edad de jubilación (65 o más años) y de 0 a 8 años de estudio, con resto de áreas urbanas y área rural, actividad primaria y Provincia de Arauco, aunque la Edad y Años de Estudio disminuyeron su aporte a las dimensiones. Bajo el eje horizontal, continúa una relación a pesar de disminuir su aporte a las dimensiones, rangos etarios jóvenes hasta 34 años, hogares con 3 a 6 o más personas, individuos soltero(a) y ahora conviviente, igualmente junto a Condición de desocupado. Además, próximo a desocupado y al grupo cercano, se encuentra actividad secundaria y Provincia del Biobío (teniendo en estas dimensiones su mejor representación) e individuos de Sexo hombre, representando una nueva relación aunque más débil, en el mismo cuadrante se encuentra de 9 a 12 años de estudio y Grupo de Ocupación 2, indicando una relación más débil con la Condición de desocupado, dada la poca explicación de la inercia acumulada corregida de estas dos dimensiones en estudio.

Luego del análisis bajo 3 mapas bidimensionales de las combinaciones de los 3 primeros factores que más explican la inercia total ajustada en Agosto para la Región del Biobío, las conclusiones varían en intensidad o importancia, pasando por algunas características demográficas, sociales y/o profesionales. Se concluye que las principales relaciones globales o de mayor importancia (ya que posee inercias originales fuertes y lógicamente mayor inercia

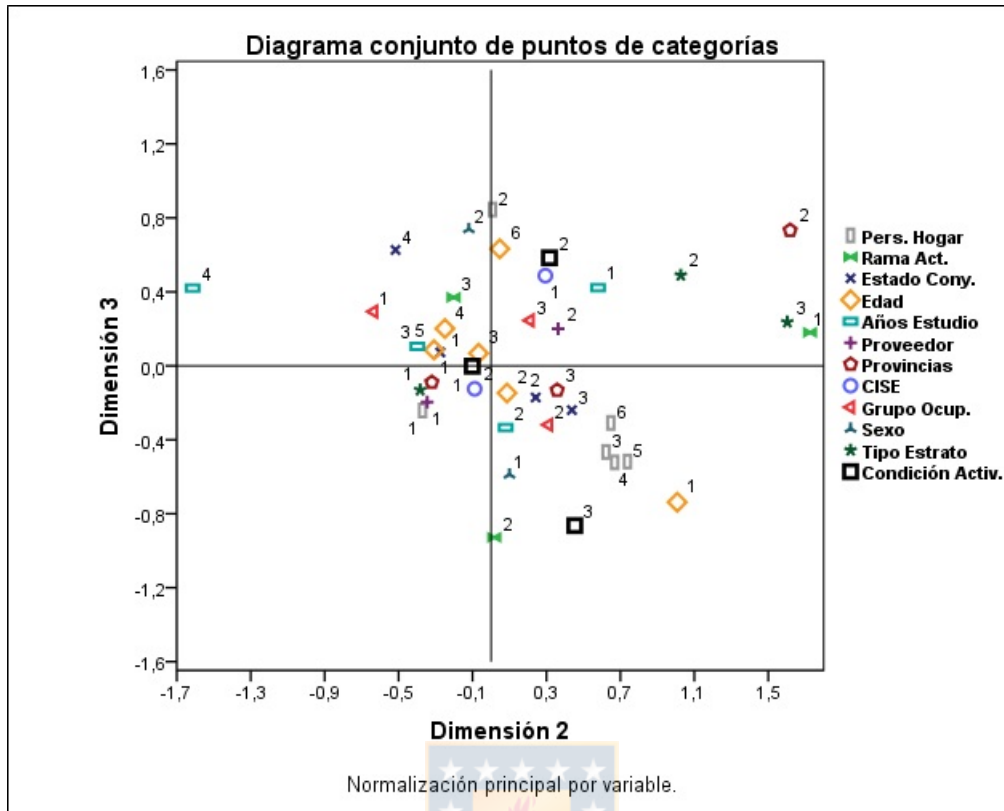


Figura 5.22: ACM de todas las variables, D2 vs D3 Agosto.

explicada acumulada corregida), son rangos etarios mayor a 35 años (el rango de 35 a 44 años unido a Ocupados), individuos casado(a) y viudo(a) o divorciado(a) que son proveedor y pueden vivir solos en el hogar, con hasta 12 años de estudio; también se identificó una relación de actividad primaria (agricultura, ganadería, casa y silvicultura, o pesca) con área rural y Provincia de Arauco, próximo de igual forma a RAU e individuos sin estudio o hasta 8 años de estudio. Además, una relación de individuos que poseen estudios superiores, Grupo 1 de Ocupación (directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleados de oficina), junto a Provincia de Concepción y de Estrato ciudades. Por último, la relación de mayor interés en este trabajo, corresponde a individuos principalmente de 15 a 24 años y de 25 a 34 años, en hogares de 3 a 6 o más personas, soltero(a), que no son proveedor junto a la Condición de desocupado. Condición de Actividad no tiene una gran Medida de Discriminación en los 3 factores, sin embargo, en la combinación de D1 vs D3 tiene su representación más fuerte y, confirma estar próximo en el grupo descrito en el mismo cuadrante, inclusive en la representación de menor importancia D2 vs D3 dada su baja explicación de la inercia ajustada.

Al analizar la D1 vs D3, se presentan próximas la Provincia de Concepción, el Estrato ciudades y los individuos con Empleo dependiente (en su representación más fuerte), en el mismo cuadrante del grupo relacionado con Desocupado, pero muy cercanas al origen dado que poseen mayor peso (tienen modalidades que no están equilibradas), implicando una relación

más específica o un fenómeno más puntual con el grupo relacionado con la Condición de Desocupado, especialmente la CISE dependiente que tiene su mejor medida de discriminación.

En el análisis de D2 vs D3, se logra apreciar una relación de menor importancia (dado que solo representan un 28,12 % de la inercia acumulada corregida) o, quizás, relativa a pocos individuos. Junto al grupo relacionado con Desocupado (individuos de 15 a 34 años, en hogares de 3 a 6 o más personas y soltero) se ubica actividad secundaria y Provincia del Biobío (en su representación más fuerte), de Sexo hombre, individuos de 9 a 12 años de estudio y Grupo 2 de Ocupación, indicando una relación más débil con el grupo de Condición de Desocupado.

El siguiente Análisis de Correspondencia Múltiple, se realizó al mes central de Noviembre, en primera instancia se realizó la Corrección de Benzecri para las inercias principales, así poder interpretar una inercia explicada corregida. Se observa en la **Tabla 5.12** las inercias principales retenidas (provenientes de cálculos de una TDC), las inercias ajustadas o corregidas junto al porcentaje de inercia explicado y acumulado.

Dimensión	Inercia TDC	Corrección Benzecri	% Inercia explicada	% Inercia acumulada
1	0,24990	0,0330199	63,33314	63,33314
2	0,17454	0,0098997	18,98802	82,32115
3	0,14376	0,0043448	8,33346	90,65461
4	0,12912	0,0024953	4,78615	95,44076
5	0,11242	0,0010072	1,93175	97,37251
6	0,10826	0,0007396	1,41853	98,79104
7	0,10165	0,0003993	0,76589	99,55693
8	0,09443	0,0001466	0,28122	99,83814
9	0,09006	0,0000538	0,10328	99,94142
10	0,08705	0,0000164	0,03147	99,97289
11	0,08663	0,0000130	0,02486	99,99775
12	0,08421	0,0000009	0,00176	99,99951
13	0,08380	0,0000003	0,00049	100,000
Suma	1,545842	0,0521368	100,000	

Tabla 5.12: Corrección de Benzecri para inercias de ACM en noviembre.

La representación de la proyección de este nuevo ACM muestra que la primera dimensión explica un 63,33 % de la inercia total ajustada, la segunda dimensión explica un 18,99 % de la inercia ajustada o corregida, y por último, la dimensión 3 explica un 8,33 % de toda la inercia ajustada. Las 3 dimensiones acumulan una inercia ajustada que explica el 90,65 %, al igual que ACM anterior. Los análisis se realizaron de manera equivalente que para mes central de Agosto, se analizó combinaciones de a pares de las 3 primeras dimensiones.

En relación a las Medidas de Discriminación, se puede observar los valores de ellas en la **Tabla 5.13**, un breve análisis y se concluye que poseen idéntica relación con los factores de las

medidas de discriminación del ACM en mes de Agosto. Nuevamente, el Número de Personas en el Hogar, Edad, Años de Estudio, Estado Conyugal y Proveedor están muy relacionadas al factor 1. Luego, Provincia y Tipo Estrato están más relacionadas al factor 2. Por último, la variable Sexo y Rama de Actividad tienen mayor relación con el factor 3. Las variables Grupo de Ocupación, CISE y Condición de Actividad en general tienen una baja relación con los factores, pero su relación más fuerte se encuentra con el factor 2, factor 1 y factor 3 respectivamente.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Provincia	0,0449	0,4330	0,1145
Tipo Estrato	0,0792	0,5316	0,1042
Personas en el Hogar	0,5941	0,1410	0,2722
Edad	0,5106	0,1233	0,0990
Sexo	0,1159	0,0156	0,4203
Años de Estudio	0,2910	0,1790	0,0805
Estado Conyugal	0,4050	0,1093	0,0563
Proveedor	0,5240	0,0858	0,0226
Grupo de Ocupación	0,1381	0,1709	0,1382
Rama de Actividad	0,1451	0,2729	0,3308
CISE	0,0805	0,0063	0,0066
Condición Actividad	0,0705	0,0258	0,0800
Promedio (Inercia)	0,2499	0,1745	0,1438

Tabla 5.13: Medidas de discriminación de ACM en noviembre.

Primer análisis, la dimensión 1 (D1) vs la dimensión 2 (D2), es el de mayor importancia por la cantidad de inercia explicada ajustada, en este plano se explica un 82,32 %. Se observa la **Figura 5.23** de representación de las variables en ACM. El gráfico corresponde a las medidas de discriminación para la D1 y la D2, como se mencionó y observó de la Tabla 5.13, para este mes central las variables numéricas categorizadas más Estado Conyugal y Proveedor estructuran el primer factor. El lugar de procedencia del individuo estructura el factor 2, junto a Grupo de Ocupación con menor fuerza y para este análisis Rama de Actividad. Las variables Sexo, CISE y Condición de Actividad poseen baja relación con las 2 dimensiones, al igual que mes central de Agosto.

Los diagramas conjuntos de los puntos de modalidades de los 3 ACM para el mes central de Noviembre mantendrán los mismos formatos (símbolo y color) para cada variable. La **Figura 5.24** representa las 2 primeras dimensiones, tras una relación de las variables con las dimensiones idénticas que mes anterior, se observa rápidamente que los posicionamientos de cada modalidad son prácticamente idénticos al ACM del mes anterior, se aprecian muy leves variaciones de algunas modalidades. Luego, sus modalidades alejadas del origen (por ende de poco peso o frecuencia) son, por ejemplo: de 19 o más años de estudio, actividad primaria y, también podría ser área rural. Por el contrario, las modalidades más cercanas al origen (de mayor peso o frecuencia) son, por ejemplo: ocupado y en este análisis otros ocupados (luego se hará una mención), de 34 a 44 años, de 9 a 12 años de estudio y el Empleo (CISE)

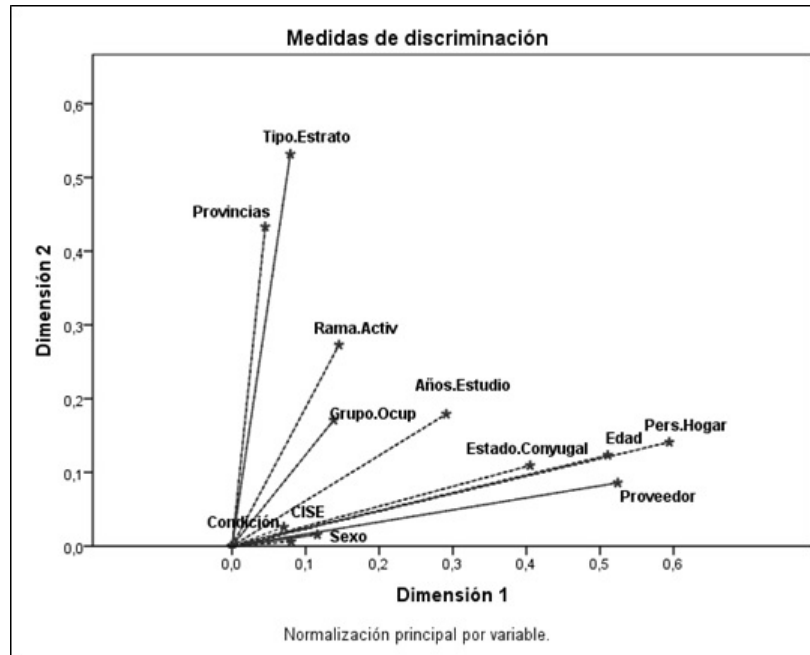


Figura 5.23: Representación de las variables en ACM, D1 vs D2 noviembre.

dependiente. Todas idénticas que en mes central de Agosto, a excepción de otros ocupados.

En relación a las proximidades de una misma variable, se mantiene la semejanza entre individuos con hogares de 3 a 6 o más personas, individuos de 45 a 54 años con los de 55 a 64 años, individuos casado(a) y viudo(a), separado(a), anulado(a) o divorciado(a), e individuos de Grupo de Ocupación 2 (Vendedores, agricultores, pescadores, operarios de construcción, trabajadores de artes, alimentos y otros) y el grupo 3 (Asistentes, peones, ayudantes, vendedores ambulantes, recolectores, otros no identificados). Sin embargo, en mes de noviembre se aprecia una semejanza más, la Condición de ocupado con la Condición de otros ocupados, ambas próximas entre sí y al origen del plano, lo que implica una semejanza entre el grupo de individuos ocupado y el grupo de individuos otros ocupados, dado que poseen un comportamiento similar en las otras variables.

Las variables más relacionadas al factor 1 se comportan de manera idéntica que en mes de Agosto. Edad se encuentra estructurada en el factor 1, de derecha a izquierda con los rangos etarios más jóvenes hasta la edad de jubilación; de forma equivalente, Número de Personas en el Hogar va de izquierda (bajo el eje y 1 persona) a derecha, aquí acumula de 3 a 6 o más personas en el hogar; Estado Conyugal ubica a soltero(a) en lado derecho y el resto a la izquierda (conviviente muy cerca del origen); Proveedor ubica sus modalidades de forma simétrica; Años de Estudio se ubica de izquierda a derecha, de 0 a 18 años de estudio y 19 o más años de estudio se ubica muy alejada del origen bajo el eje. A grandes rasgos la estructuración del factor 1 es análoga al mes de Agosto, salvo leves cambios de posicionamiento de algunas modalidades.

En efecto, sobre el factor 2, la semejanza con el mes de Agosto está a la vista, Provincia y Tipo Estrato se ubican en la extensión del eje vertical, ubicando Concepción y ciudades bajo el eje y el resto por sobre el eje; Grupo de Ocupación 1 y actividad terciaria se ubican en cuadrante inferior derecho, el resto se ubican en cuadrante superior izquierdo.

Nuevamente se distinguen a izquierda del plano, individuos de rangos etarios mayores a 35 años, casado(a) y viudo(a) o divorciado(a), que son proveedores y pueden vivir solos en el hogar, con hasta 12 años de estudio; por sobre el eje se relaciona la actividad primaria (agricultura, ganadería, caza y silvicultura o pesca) con área rural y RAU, Provincia de Arauco, individuos sin estudio y hasta 8 años de estudio. En el mismo lado izquierdo del plano, las Condiciones de ocupado y otros ocupados, aunque muy cercanas al origen. Un leve cambio sufren los individuos de 19 o más años de estudio, ubicada muy alejada del origen y en lado izquierdo, el mismo cuadrante que individuos mayores a 35 años y Condición de ocupado.

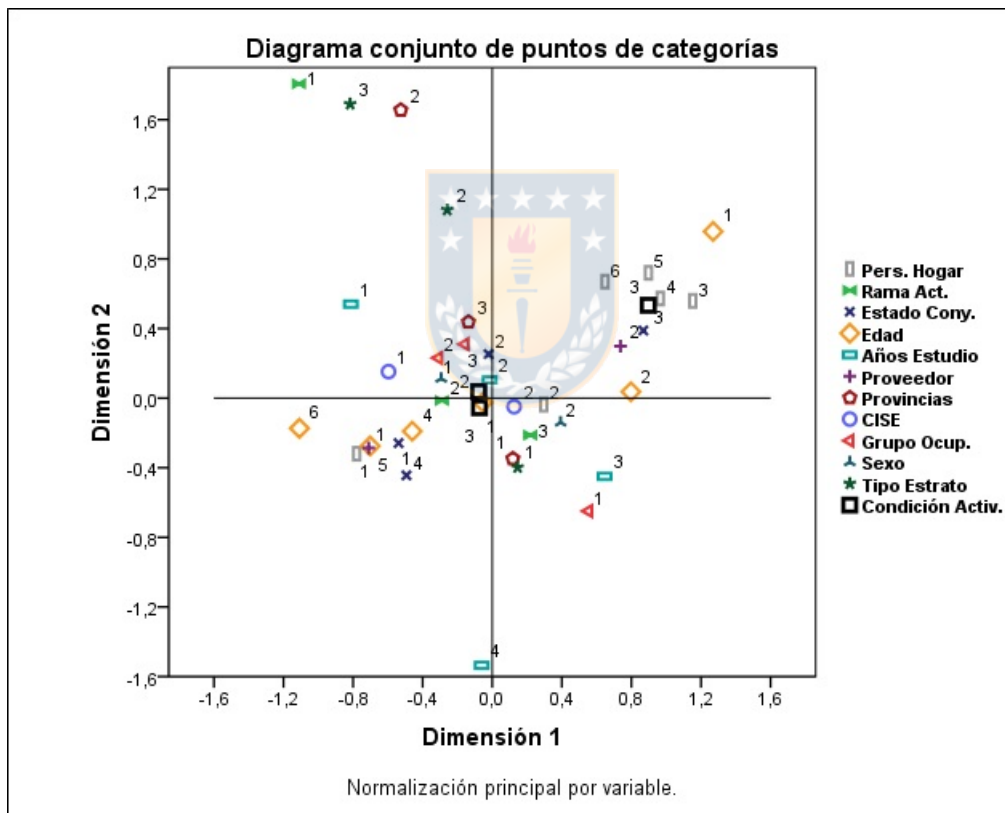


Figura 5.24: ACM de todas las variables, D1 vs D2 noviembre.

En el lado opuesto del plano, lado positivo del factor 1 pero bajo el eje, se ubica Provincia de Concepción unido a ciudades, próximos a Grupo de Ocupación 1 (directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio, y empleados de oficina), de 13 a 18 Años de Estudio y actividad terciaria (comercio al por mayor y menor, hoteles y restaurantes, transporte, actividades inmobiliarias, administración pública, enseñanza, entre otros), idéntico que mes de Agosto, salvo los individuos de más Años de Estudio (19 o más).

Esta última y más relevante relación se vuelve a repetir, en cuadrante superior derecho se ubica rangos etarios jóvenes hasta 34 años, que viven de 3 a 6 o más personas en el hogar, soltero(a) y no proveedor, junto a la Condición de desocupado que a pesar de su baja relación (pero más fuerte en el factor 1) la ubica perfectamente próxima a ellos. Idéntica conclusión con ACM de Agosto.

Segundo análisis, la dimensión 1 (D1) vs la dimensión 3 (D3), la inercia ajustada explicada en este plano es de 71,66 %, disminuye por una menor inercia ajustada de la dimensión 3. Se observa la **Figura 5.25** de representación de las variables en ACM y corresponde a las medidas de discriminación para la D1 y la D3. Las variables más relacionadas al factor 1 se explicaron en primer análisis. Luego, Sexo y Rama de Actividad toman mayor relación al factor 3, pierden relación en este análisis Provincia, Tipo Estrato y algo de representación Grupo de Ocupación, situación idéntica que mes de Agosto. La variable Condición de Actividad presenta su mejor relación con las dimensiones en este análisis, aunque globalmente bajas, igual que mes de Agosto.

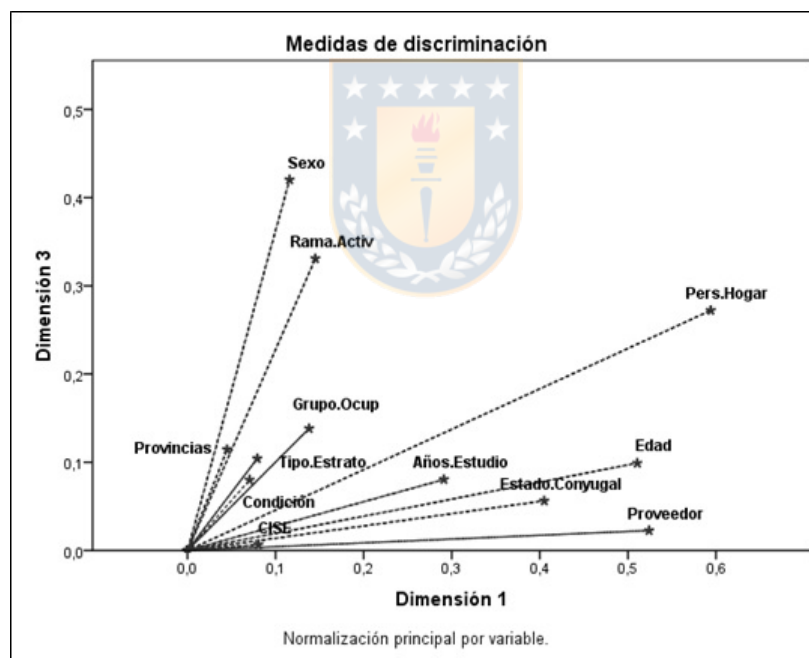


Figura 5.25: Representación de las variables en ACM, D1 vs D3 noviembre.

El estudio de la D1 vs D3 (ver **Figura 5.26**), implica un análisis análogo al realizado en mes de Agosto para las mismas dimensiones y, efectivamente se obtienen resultados en su mayoría equivalentes, en cuanto a su lejanía o cercanía al origen y a la proximidad entre modalidades de una misma variable, pero en este mes se agrega la Condición de ocupado próxima a la Condición de otros ocupados, lo que indica una semejanza en los individuos que poseen esa Condición de Actividad en el resto de las variables.

La relación de las variables con el factor 1 fueron descritas en primer análisis de Noviembre, y el posicionamiento de las modalidades es equivalente a mes de Agosto. Edad se estructura de derecha (sobre el eje) a izquierda (bajo el eje); Número de Personas en el Hogar en forma de parábola vertical, dejando sus extremos (1 persona y 3 a 6 o más personas) sobre el eje, mejora su relación en el factor 3; Años de Estudio en forma de elipse horizontal y acerca de 19 o más años de estudio (de poca frecuencia), disminuye su relación en el factor 3; Estado Conyugal ubica a soltero(a) en derecha y sobre el eje, el resto quedan en izquierda y bajo el eje; Proveedor de forma simétrica pero disminuye su relación en el factor 3, por lo que sufre un cambio de posición.

Las relaciones de las variables del factor 3 son equivalentes a los resultados del mes de Agosto. Sexo es simétrico con respecto al origen; Rama de Actividad ubica a actividad secundaria en lado positivo y la opone de las otras actividades; Número de Personas en el Hogar tiene mejor relación en factor 3 que factor 2; CISE tiene casi idéntica relación en factor 3 y 2, sin embargo las modalidades se mueven de cuadrantes; Condición de Actividad en su mejor representación deja desocupados en lado positivo del factor 3 y alejada del origen, las otras Condiciones quedan en lado negativo y muy cerca del origen y entre ellas.

Las principales relaciones obtenidas nuevamente son análogas al mes de Agosto, ya sabemos que la dimensión 1 separa individuos adultos jóvenes hasta 34 años y mayores, separa Número de Personas en el Hogar, ser o no Proveedor, ser soltero(a) de otro Estado Conyugal. Además, se relaciona a hombres con actividad secundaria (explotación de minas y canteras, industrias manufactureras, suministro de electricidad, gas y agua, y construcción), que viven solos en el hogar y son proveedor, aunque bajan su representación Grupo de Ocupación 2 y de 9 a 12 años de estudio se ubican en mismo cuadrante que grupo mencionado. Luego, se relaciona actividad primaria con Edad de jubilación, se mantiene próximo área rural a pesar que Estrato disminuye su relación con las dimensiones, también relacionados a individuos sin estudio y hasta 8 años de estudio. Estas 2 relaciones son equivalentes a las encontradas en mes de Agosto.

Además, otra relación que viene representada en la dimensión 3, y análoga que en mes de Agosto, individuos de Sexo mujer en hogar de 2 personas, próximo a actividad terciaria (comercio al por mayor y menor, hoteles y restaurantes, transporte, actividades inmobiliarias, administración pública, enseñanza, entre otros) y a Grupo de Ocupación 1, ésta última cercana a 13 a 18 años de estudio.

La Condición de Actividad en su mejor relación con las dimensiones, mantiene a individuos desocupados en el mismo cuadrante de adultos jóvenes hasta 34 años (mucho más próximo de 15 a 24 años), hogares de 3 a 6 o más personas y soltero(a), confirmando la relación del primer análisis y del mes de Agosto. Proveedor pierde relación en el factor 3 y no se encuentra próxima pero en mismo lado derecho. CISE tiene la misma baja relación en factor 2 y factor 3, pero cambia de posición y ubica a individuos dependientes en mismo cuadrante de desocupado, y en igual posición que mes de Agosto. Unidas Provincia de Concepción y Estrato ciudades, en el mismo cuadrante, pero en este análisis poseen una baja representación en las dimensiones. La ubicación de estas 3 últimas modalidades es idéntica que mes de

Agosto, y puede ser una relación más puntual o específica solo de algunos individuos, con menor importancia dada la inercia ajustada de la dimensión 3.

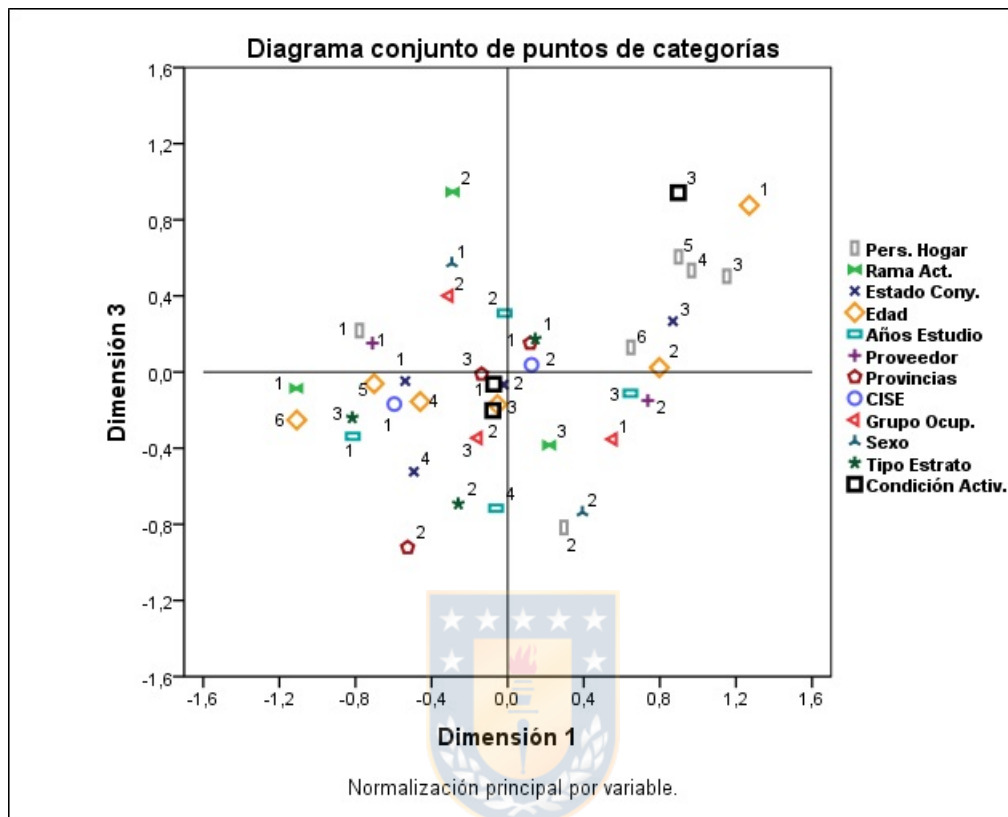


Figura 5.26: ACM de todas las variables, D1 vs D3 noviembre.

Tercer y último análisis, la dimensión 2 (D2) vs la dimensión 3 (D3), de menor importancia dado que el plano explica un 27,32% de la inercia total ajustada. Se observa la **Figura 5.27** de representación de las variables en ACM, que corresponde a las medidas de discriminación para la D2 y la D3. Las variables de mayor relación con el factor 2 son Provincia y Tipo Estrato. Las variables de mayor relación con el factor 3 son Sexo, Número de Personas en el hogar y Rama de Actividad (en su mejor relación con las dimensiones). Pierde relación con las dimensiones Edad, Estado Conyugal, Proveedor y CISE, Años de Estudio es baja su relación al factor 3. Grupo de Ocupación tiene igual relación que en análisis de D1 vs D2. La Condición de Actividad tiene su más fuerte relación en factor 3. Análisis prácticamente análogo que mes de Agosto, salvo algunas excepciones mínimas.

El diagrama conjunto de modalidades de las dimensiones D2 vs D3 se observa en la **Figura 5.28**, dónde nuevamente los principales resultados son equivalentes al mes de Agosto. Sobre las modalidades más alejadas y más próximas al origen se han estudiado en los análisis anteriores, en esta ocasión hay un mayor conjunto de modalidades muy cercanas al origen, por ejemplo: CISE que pierde relación con ambas dimensiones tiene sus modalidades centradas,

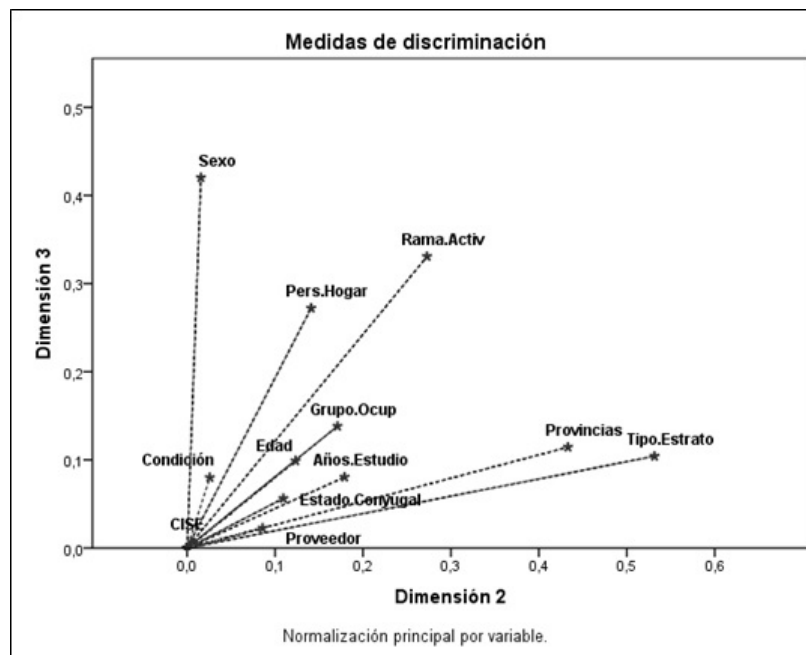


Figura 5.27: Representación de las variables en ACM, D2 vs D3 noviembre.

sin mayor poder de explicación. La proximidad de modalidades de una misma variable es equivalente en análisis anteriores, Número de Personas en el Hogar de 3 a 6 o más personas, y en mes de Noviembre las Condiciones ocupado y otros ocupados.

Las variables Edad, Estado Conyugal, Proveedor y CISE, pierden mucha relación en este análisis de las dimensiones y ya no se distingue alguna estructuración clara en los ejes. En consecuencia, Edad tiene sus modalidades de 35 años hasta años de jubilación muy próximas entre ellas. Se estructura a lo largo del eje horizontal principalmente Provincia y Tipo Estrato, con menor relación lo hace Años de Estudio de derecha a izquierda y, Grupo de Ocupación. El eje vertical del factor 3, se relaciona principalmente a Sexo, Rama de Actividad y con menor relación Número de Personas en el Hogar, que repite su estructura de parábola vertical con sus extremos en la parte superior.

Sobre el conjunto de todas las modalidades, las relaciones que se reiteran con respecto al mes de Agosto son: individuos de Concepción, ciudades (algo lógico) y que viven solos; individuos con actividad terciaria, de Sexo mujer, Grupo 1 de Ocupación y de 13 o más años de estudio; son relaciones de menor importancia. Además, la relación de individuos con actividad primaria, área rural y RAU, con Provincia de Arauco y sin estudios o hasta 8 años de estudio, con mayor importancia dado que se reitera de primer análisis (igual que Agosto), además se incluye el Grupo 3 de Ocupación (asistentes, peones, ayudantes, vendedores ambulantes, recolectores, otros no identificados), que tiene igual relación con D3 que D1.

Se confirma y reitera la relación de análisis y mes anterior, los individuos desocupados con rangos etarios jóvenes, mucho más próximo el rango de 15 a 24 años, de 3 a 6 o más Personas

en el Hogar y soltero(a). En esta ocasión, se aproxima Grupo 2 de Ocupación, Sexo hombre, de 9 a 12 años de estudio, representando una relación más débil con el grupo próximo a desocupado por la menor importancia de las inercias corregidas; actividad secundaria y Provincia del Biobío (ambas con su mejor representación) se encuentran en los límites de los cuadrantes. Resultado casi equivalente que mes de Agosto. Se aprecia más cambio de posición de modalidades, puede ser debido a fenómenos puntuales de un mes a otro o en pocos individuos que cambian sus condiciones de actividad.

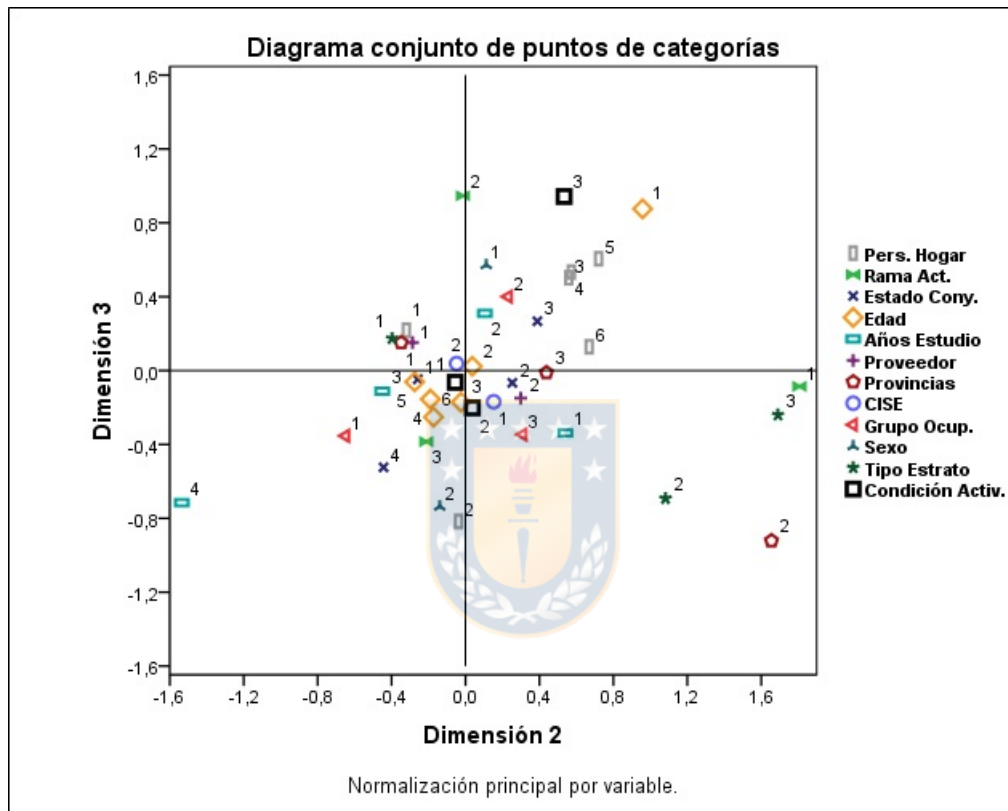


Figura 5.28: ACM de todas las variables, D2 vs D3 noviembre.

Los resultados obtenidos para el mes de Noviembre de las 3 primeras dimensiones que más explican la inercia corregida, no son extraños ni ajenos, se reiteran prácticamente todos los resultados, hay excepciones de algunas modalidades porque la variable posee una baja relación con alguna dimensión. Las principales relaciones globales en D1 vs D2 se mantienen casi idénticas, los rangos etarios mayor a 35 años (el rango de 35 a 44 años unido a ocupado) individuos casado(a) y viudo(a) o divorciado(a) que son proveedor y pueden vivir solos en el hogar, con hasta 12 años de estudio; otra relación identificada fue la de actividad primaria (agricultura, ganadería, casa y silvicultura o pesca) con área rural y Provincia de Arauco, próxima a ellas también se encuentra RAU y sin estudios o hasta 8 años de estudio. Hay 2 modalidades que cambian levemente de posición en Noviembre, una es la Edad de jubilación (de 65 o más años) ahora se aproxima más a los otros rangos etarios y proveedor que puede vivir solo, ubicado en cuadrante de ocupado, en Agosto se ubicó más próximo a individuos sin

estudios o hasta 8 años de estudio en cuadrante de otros ocupados. La otra modalidad es de 19 o más años de estudio, a pesar de su idéntica lejanía del origen en ambos meses, se ubica en cuadrante de ocupado y otros ocupados, con las cualidades ya mencionadas. Mientras que en Agosto, se ubicó en cuadrantes de desocupado.

El otro grupo de principales relaciones globales son individuos de 13 a 18 años de estudio, Grupo 1 de Ocupación (directores, gerentes, profesionales científicos, técnicos profesionales de nivel medio y empleados de oficina), junto a Provincia de Concepción y ciudades. Y de forma idéntica que en Agosto, se relaciona individuos de 15 a 24 años (más próximo) y de 25 a 35 años, de 3 a 6 o más personas en el hogar, soltero(a) y no proveedor, junto a la Condición de desocupado.

El segundo análisis de D1 vs D3, es equivalente a mes de Agosto, la modalidad no proveedor se aparta del grupo relacionado con desocupado, ya que pierde relación con el nuevo factor. El individuo con Empleo dependiente en CISE, se ubica en mismo cuadrante del grupo de relación con desocupado (aunque siempre mantiene una baja relación con los factores, esta combinación es la de mayor representación), al igual que Provincia de Concepción y Estrato ciudades, que se ubican en mismo cuadrante (ambas próximas a dependiente) pero su relación con los factores es la menor de todas, se explicaría una relación más específica o un fenómeno puntual con la Condición de desocupado.

El último análisis de las dimensiones de menor importancia, muestra firmemente la Condición de desocupado con individuos de 15 a 24 años principalmente, hogares de 3 a 6 o más personas y soltero(a). Al igual que Agosto, se aproxima a ellos Grupo 2 de Ocupación, Sexo hombre y de 9 a 12 años de estudio, dejando en evidencia una relación más débil o de menor importancia, por situación más específica y probablemente relativa a pocos individuos. En comparación con Agosto, actividad secundaria y Provincia del Biobío se ubican levemente en el borde del cuadrante del grupo próximo a desocupado, evidenciando un cambio de posición en el mes de Noviembre. Otro cambio, es para el individuo conviviente que en mes de Agosto se dejó ver próximo al grupo de desocupado (aunque con poca relación con los factores), mientras que en noviembre no se observó proximidad. Estos cambios son muy específicos, puede ser parte de las variaciones estacionales de los meses en estudio o condiciones propias de los individuos, que aleja la actividad secundaria y la Provincia del Biobío del grupo relacionado con desocupado, además no es un cambio global, ocurre en las dimensiones de menor importancia dada su representación de la inercia corregida.

5.5. Análisis provincial mediante ACM.

Se presentan los principales resultados obtenidos en el Análisis de Correspondencia Múltiple para cada una de las Provincias de la Región, Concepción, Arauco y Biobío, a modo de comparación estarán meses centrales de Agosto y Noviembre en estudio.

El conjunto de variables a utilizar es el mismo con el que se realizó los ACM para la Región en ambos meses. Como ahora habrá un ACM por Provincia, las variables serán 11 en total, incluida la variable de interés Condición de Actividad. El objetivo es analizar y comparar la estructura de relación existente en cada Provincia de la Región en los mismos 2 meses centrales consecutivos. En cada ACM se analizó en las 3 primeras dimensiones, dado que con la corrección de Benzecri se obtuvo sobre un 90 % de inercia corregida explicada en cada ACM. La tabla de inercias con la Corrección de Benzecri, la tabla de Medidas de Discriminación y el gráfico de ACM conjunto, para cada Provincia y mes se pueden encontrar en el **ANEXO** del trabajo.

La Provincia de Concepción en el mes de Agosto para las primeras 2 dimensiones, presenta una proximidad en individuos de 15 a 24 años, hogares de 3 a 6 o más personas, soltero(a) y no proveedor, junto a la Condición de desocupado (llamaremos grupo de relación global o principal), las primeras variables con una alta relación con el factor 1 y Condición de Actividad en general con una baja relación con las dimensiones en estudio. Esta estructura se ve como una relación habitual en los análisis, sin embargo, en esta ocasión aparece el Estado de conviviente en mismo cuadrante del grupo relacionado. El análisis de la dimensión 1 vs la dimensión 3, mantiene la relación principal del grupo próximo a desocupado, inclusive reitera la lejanía de no proveedor como en los análisis de la Región y, mantiene el Estado conviviente del análisis anterior en el mismo cuadrante (en su mejor relación con los factores). Además, la CISE dependiente se ubica en el mismo cuadrante de interés en su mejor relación con los factores (aunque relativamente baja). Hay otras categorías que se ubican en el cuadrante de interés, como individuos de Estrato ciudades (igual que análisis de región), de 13 a 18 años de estudio y Grupo 1 de Ocupación, pero las 3 variables tienen su peor relación con los factores. El análisis de las dimensiones de menor inercia (D2 vs D3), confirma la relación global o principal en la Provincia de Concepción, sin individuos no proveedor e incluyendo conviviente, a pesar que las variables en cuestión disminuyan su relación con estos factores y Condición de Actividad posee su mayor relación en esta combinación de factores. Otras categorías próximas a este grupo, evidenciando una relación más débil por la importancia de sus inercias, son los individuos de actividad secundaria (teniendo su mejor relación con los factores), de Sexo hombre, de Grupo 2 de Ocupación y de 9 a 12 años de estudio. Estas 4 categorías, son equivalentes a las encontradas en el análisis de la Región en el mismo mes en estudio.

El mes de noviembre para las 2 primeras dimensiones, tiene idéntica relación principal y se observa que el Estado conviviente se ubica en el mismo cuadrante pero mucho más cercano al origen que al grupo de relación. El análisis de la dimensión 1 vs la dimensión 3 tiene equivalente relación que mes anterior, dado que aleja a individuos no proveedor, mantiene el Estado conviviente (pero cercano al origen) e ingresa al cuadrante individuos dependiente (como en todos los análisis realizados en D1 vs D3) bajo su mejor relación con las dimensiones, pero ubicado cerca del origen por su alta frecuencia. Se destaca que cambia de cuadrante la modalidad 6 o más personas en el hogar, pero en mismo lado del grupo de relación principal, caso similar con individuos de 15 a 24 años aunque se justifica con su poca relación en factor 3. Igual que mes de agosto, otras categorías se ubican en el cuadrante de

interés como individuos de ciudades, de 13 a 18 años de estudio y Grupo 1 de Ocupación, e igual, las 3 variables en cuestión tienen su peor relación con los factores. Para el análisis de la dimensión 2 vs la dimensión 3, la relación principal se mantiene aunque se ve afectada por la baja relación de las variables con los factores, no proveedor y 6 o más personas en el hogar se alejan del grupo de relación, conviviente se mantiene más próxima al origen que a desocupado. Por otra parte, las nuevas modalidades que se aproximan al grupo de relación principal son actividad secundaria, Sexo hombre y Grupo 2 de Ocupación, las 3 variables en cuestión en su mayor relación con los factores. Este grupo representan una relación más débil por la menor inercia explicada, son las mismas modalidades encontradas en mes de agosto para la Provincia de Concepción, a excepción de 9 a 12 años de estudio que se ubica levemente fuera del cuadrante del grupo de relación principal, por ende hace compleja una interpretación adecuada para mes de noviembre.

La Provincia de Arauco en el mes de Agosto para las primeras 2 dimensiones, mantiene el grupo de relación global o principal, pero ahora incluye individuos de 25 a 34 años y de 9 a 12 años de estudio, por primera vez en los análisis. El Estado conviviente vuelve a tener idéntica ubicación, muy próxima al origen y en mismo cuadrante del grupo de relación principal. Se destaca que, la Provincia de Arauco no posee el Estrato ciudades, solo existen los estratos RAU y área rural. El análisis de la dimensión 1 vs la dimensión 3 presenta mayor variabilidad, se alejan soltero(a) y no proveedor ya que su relación con el factor 3 disminuye, por el contrario, se aleja hogares con 3 personas, en su mejor representación. La Provincia de Arauco presenta mayor relación en individuos de 15 a 34 años, en hogares de 4 a 6 o más personas, de 9 a 12 años de estudio, que están desocupado. El Empleo dependiente se aproxima a desocupado y al grupo en general, en su mejor relación con los factores, situación idéntica que Provincia de Concepción y análisis regional. Luego, otras categorías se aproximan al grupo como RAU y Grupo 1 de Ocupación, pero con bajas relaciones con las dimensiones. En el estudio de la dimensión 2 vs la dimensión 3, la relación que se mantiene en Arauco principalmente es de individuos de 15 a 34 años, hogares de 4 a 6 o más personas, de 9 a 12 años de estudio, con la Condición de desocupado. El Sexo hombre, se ubica próximo al grupo de relación, como ha sido tónica en Concepción y a nivel regional, teniendo una relación más débil o puntual dada la cantidad de inercia explicada. El comportamiento de conviviente se mantuvo en todas las combinaciones de dimensiones, en el borde del cuadrante de relación principal. Además, próximo al grupo de relación mencionado, se ubica actividad primaria, sin embargo posee la más baja relación con los factores; actividad secundaria se ubica en el límite de ingresar al cuadrante (como se ha visto en Concepción y en la Región), pero por su baja relación con el factor no tiene mejor posición en el mapa.

El mes de noviembre para la dimensión 1 vs la dimensión 2, presenta idéntica relación principal que mes de agosto, los individuos más relacionados con Condición de desocupado son de 15 a 34 años, con hogares de 3 a 6 o más personas, soltero(a), no proveedor y con 9 a 12 años de estudio. La dimensión 1 vs la dimensión 3, nuevamente presenta mayor variabilidad, Número de Personas en el Hogar en su mejor representación por los factores solo deja próximo a desocupado hogares de 6 o más personas, aunque hogares de 3 a 5 personas se ubican en el mismo lado que desocupado, cambian de cuadrante. Luego, se mantiene in-

dividuos de 15 a 24 años (bajando su relación en factor 3), no proveedor y con 9 a 12 años de estudio, próximos a desocupado. El Empleo dependiente no se aproxima a Condición de desocupado, a diferencia del mes de agosto, de la Provincia de Concepción y a modo regional, por única vez en los análisis. Otras categorías se aproximan a desocupado como Sexo mujer (única vez en los análisis) y actividad terciaria, pero con sus más bajas relaciones en las dimensiones, esto no se observa en agosto. El análisis de las dimensiones de menor inercia (D2 vs D3), mantiene relación de individuos de 15 a 24 años, hogares de 6 o más persona, de 9 a 12 años de estudio y no proveedor, junto a desocupado, a pesar que se encuentran en sus más bajas relaciones con las dimensiones. Otras categorías próximas a desocupado son Estrato rural y Grupo 2 de Ocupación, que para Arauco en agosto no se encontraron, éstas responden a una relación más débil con desocupado, dada por algunos individuos específicos o situación puntual. A diferencia de agosto, el Sexo y la Rama de Actividad, no presentan alguna relación específica con la Condición de desocupado.

La Provincia de Arauco presentó más dinamismo en sus relaciones, en general se obtuvo una relación global o principal concordante con el análisis regional, pero se incluye de manera concisa los individuos de 9 a 12 años de estudio en ambos meses. También, se observó una tendencia de relación con Condición de desocupado a hogares con 4 o más persona, inclusive en noviembre hogares con 6 o más personas, bajo su mejor representación.

La Provincia del Biobío en el mes de Agosto para el análisis de las primeras 2 dimensiones, presenta una relación principal en individuos de 15 a 24 años, hogares de 3 a 6 o más personas, soltero(a) o conviviente, y no proveedor, junto a Condición de desocupado. La dimensión 1 vs la dimensión 3 mantiene casi la misma relación principal, se aleja no proveedor (como es de costumbre en el análisis D1 vs D3) y algunas modalidades de personas en el hogar producto de que disminuye la relación con factor 3. Condición de Actividad presenta su peor representación en las dimensiones, mantiene sus modalidades muy cercanas al origen, pero la Condición de desocupado se mantiene en el cuadrante de la relación principal. El Empleo dependiente y de 13 a 18 años de estudio, se ubican en cuadrante y próximos del grupo de relación principal, en sus mejores relaciones con los factores. Por otra parte, el Estrato ciudades y el Grupo 1 de Ocupación, también se ubican próximos a desocupado o en el cuadrante del grupo de relación principal, pero tienen una baja relación con los factores. En las dimensiones de menor inercia (D2 vs D3), se reitera la relación principal sobre individuos en dichos rangos etarios y hogares, con Estado soltero(a) o conviviente, junto a desocupado. Los individuos de Estrato RAU, actividad secundaria, Sexo hombre, de 9 a 12 años de estudio y Grupo 2 de Ocupación, se ubican próximas y en mismo cuadrante que relación principal con desocupado, con relaciones similares con los factores, por tanto se evidencia una relación más débil por la menor inercia explicada, y representada por pocos individuos o situaciones puntuales de estacionalidad.

En el mes de noviembre para la dimensión 1 vs la dimensión 2, resulta la misma relación principal que mes de agosto, el único cambio es que individuo conviviente se aleja de desocupado y del grupo para quedar cercana a la Condición de otros ocupados, en agosto conviviente tuvo proximidad con el grupo de relación principal en Biobío. El análisis de la

dimensión 1 vs la dimensión 3 mantiene la tendencia general de alejar a no proveedor del grupo principal, el resto continua próximo entre sí, pero sin el individuo conviviente este mes. Luego, como es de costumbre en este análisis, ingresa al cuadrante del grupo principal de relación el Empleo dependiente (muy cerca del origen, bajo su mejor relación con los factores), reiterando la relación encontrada al analizar la dimensión 3. A diferencia de agosto, Años de Estudio en su mejor representación no logra ingresar ninguna modalidad al cuadrante o grupo relacionado, aunque de 13 a 18 años de estudio y de 9 a 12 años de estudio se ubican en los márgenes de ingresar. El análisis de las dimensiones de menor inercia muestra qué, inclusive cuando las variables disminuyen su relación con los factores, la relación global o principal se mantiene pero sin incluir a no proveedor (en su más baja relación con los factores) y solo con individuo soltero en este mes (en agosto, también se relacionaba con conviviente). Luego, los individuos de Estrato RAU, Sexo hombre, de 9 a 12 años de estudio y Grupo 2 de Ocupación, ingresan al cuadrante de la relación global y aparte se mantienen muy próximas entre ellas, lo que indica una relación más débil de pocos individuos o situación específica. Rama de Actividad bajo su mejor relación con las dimensiones, no se aproxima de forma notoria a alguna Condición de Actividad. En agosto, se encuentran las mismas modalidades en este análisis, reiterando los individuos con una relación más débil, excepto actividad secundaria que en noviembre no resultó con una proximidad interesante de interpretar.

La Provincia del Biobío se mostró con relaciones similares en los meses de estudio, mantuvo la relación principal incluyendo a conviviente en agosto, para que en noviembre no se encontrara próximo. Sobre la relación de menor interés y más débil, también fue constante en los meses de estudio, solo actividad secundaria se vio afectada por la no relación y poca proximidad con desocupado y el grupo de relación principal.

Capítulo 6

Conclusión, discusión y proyecciones.

El análisis descriptivo que se realizó a las variables de trabajo en ambos meses centrales para la Región del Biobío, fue a modo de describir las modalidades involucradas, las mayores y menores frecuencias de registros de cada variable y las diferencias entre meses centrales. Esta diferencia también se puede ver como un cambio de estacionalidad, ya que al tener registros de mes central de agosto, involucra encuestados en mes de Julio, Agosto y Septiembre, observando que son meses en estación de invierno; al tener registros de mes central de noviembre, involucra encuestados en mes de Octubre, Noviembre y Diciembre, entendiéndose como meses en estación de primavera.

En primera instancia se destacó que el índice de desempleo para agosto fue de un 7,9% y para noviembre fue de un 6,9%, el mayor y menor índice de desempleo de los 4 trimestres móviles con que se obtienen los promedios anuales. La diferencia fue de 1%. Mientras que, los datos empíricos arrojaron una tasa bruta de desocupados de 8,8% en agosto y de un 8,1% en noviembre, confirmando el mayor porcentaje en agosto y evidenciado una diferencia de 0,7%. Se debe tener en cuenta que se trabajó con la base de datos en bruto, sin utilizar el factor de expansión para los análisis. Luego, la Condición de otros ocupados fue mayor en agosto por 2,1% y, la Condición de ocupado fue mayor en noviembre por 2,82%, siendo de 80,85% vs un 78,03% en agosto.

A modo general, las modalidades de las variables en estudio tuvieron registros muy variados, categorías de alta o poca frecuencia en la mayoría de las variables, salvo Sexo que tuvo registros de hombre/mujer de 56%/44% aproximado en ambos meses y, proveedor tuvo 50% de registros si era proveedor o no (la más uniforme de todas), también en ambos meses de estudio. Unas variables fueron proporcionales a la población como Provincia y Tipo de Estrato, sus mayores registros fueron en Concepción y ciudades respectivamente (alrededor de 73% y de 76% en ambos meses), sin embargo, responde a la naturaleza del muestreo realizado para la medición de la ENE.

Algunas variables nos mostraron cualidades de la población como: Estado Conyugal con un 42% aproximado de registros casado(a) y un 33% aproximado de soltero(a), Número de Personas en el Hogar con un 45% de individuos aproximado que viven solos, un 27% aproximado de hogares con 2 personas y un 15% aproximado de hogares con 3 personas;

Edad de 25 a 34 años con un 21 % aproximado de los registros, al igual que de 35 a 44 años, casi un 25 % de los individuos con Edad de 45 a 54 años y destaca un 5 % de registros de 65 o más años; Años de Estudio de 9 a 12 años en un 42 % aproximado de los registros y un 32 % aproximado de 13 a 18 años de estudio. Todos los porcentajes de las modalidades fueron prácticamente equivalentes en ambos meses de estudio, las variaciones fueron mínimas.

Otro grupo de variables nos mostró aspectos económicos de la población como: Grupo de Ocupación con un 45 % de registros aproximado en Grupo 2 (vendedores, agricultores, pescadores, operarios de construcción, trabajadores de artes, alimentos y otros) y un 25 % de registros aproximado en Grupo 3 (trabajadores no calificados u otros no identificados); Rama de Actividad económica con un 7 % de registros aproximado en actividad primaria y un 66 % aproximado de individuos en actividad terciaria; Situación del Empleo (CISE) con un casi 20 % de empleos independientes y un 80 % aproximado de individuos con empleo como dependientes. Los porcentajes son válidos para ambos meses en estudio, dado la poca variación mencionada, las diferencias de las modalidades no fueron mayores a 2 % del mes de agosto al mes de noviembre.

Bajo este análisis descriptivo y general de las variables sociodemográficas y económicas, se logro caracterizar a la Fuerza de Trabajo de la Región del Biobío, y se observó o demostró que las muestras (individuos) de la población en estudio para los meses centrales son muy homogéneas, no se encuentran grandes diferencias porcentuales de un mes a otro en los individuos encuestados. A pesar que hay un cambio de estación, las cualidades de la población o los aspectos económicos no varían de agosto a noviembre. Este efecto de no variación de los registros de individuos en los meses a comparar, puede ser respuesta al no cambio del muestreo que realiza INE, dado que para área urbana, restos de áreas urbana y área rural los cambios de muestra ocurren luego de 18, 27 y 36 meses respectivamente. También, puede ser debido a que la población en general no sufrió de mayores cambios en aspecto sociodemográfico o laboral (económico). Sin embargo, la Condición de Actividad tuvo variación de un 0,7 % en desocupado y, la tasa de desempleo tuvo una variación de 1 %, siendo siempre mayor en mes de Agosto con respecto a Noviembre.

El análisis de las Tablas de Contingencia de cada variable por la Condición de Actividad nos abrió la visión a resultados interesantes, se observó mediante perfiles de columna la distribución que poseía cada variable de trabajo según cada Condición de Actividad, se comparó principalmente la distribución de cada variable dado los registros de individuo desocupado en mes de agosto con mes de noviembre. Se destaca en primer lugar que, todas las tablas de contingencia fueron analizadas mediante la prueba estadística Chi-Cuadrado de Independencia, las cuales en su totalidad resultaron con un valor p significativo ($< 0,05$), lo que significó que cada variable estaba relacionada con la Condición de Actividad, en ambos meses de estudio.

Una segunda conclusión, sobre la comparación de los resultados de agosto a noviembre, nos muestra que los mayores registros de desocupados en cada una de las variables fueron los mismos en ambos meses, salvo una leve variación en el Número de Personas en el Hogar y Edad. El resto de variables mantuvo la mayoría en la misma modalidad de los registros de

individuos desocupados.

En este análisis cruzado, algunas variables siguieron el comportamiento general, ya que mantuvieron una mayoría de registros desocupados en las mismas modalidades donde tuvieron mayores frecuencias de registro general, como: Provincia y Tipo Estrato que tuvieron una mayor cantidad de desocupado en Concepción (alrededor de 78 %) y una mayor cantidad de desocupado de ciudades (alrededor de 82 %), la variable Sexo que su mayor porcentaje de desocupado es hombre (58 % aproximado), Grupo de Ocupación que su mayor porcentaje de desocupado son de Grupo 2 (casi 50 % en agosto y alrededor de 45 % en noviembre), de igual forma Rama de Actividad tuvo el mayor porcentaje de desocupado en actividad terciaria (52 % y 57 % aproximado en agosto y noviembre), CISE tuvo su mayor registro de desocupado en empleo dependiente (alrededor de 95 %) y, la variable Años de Estudio mostró que los mayores registros de desocupados tienen de 9 a 12 años de estudio (alrededor de 49 %).

Otro grupo de variables, tuvo discrepancias en las frecuencias relativas de los niveles de desocupado en cierta variable con la mayoría de registros de la misma variable vista en modo marginal, como: Estado Conyugal que mostró una mayoría de registros casado(a) (42 % aproximado) en ambos meses, mientras que la mayoría de desocupados son soltero con un 47 % aproximado en agosto y aumenta a un 56 % aproximado en noviembre; la variable Proveedor tuvo una igualdad de individuos que era proveedor y que no era proveedor (50 % aproximado), mientras que la mayoría de los desocupados son no proveedor con un 73 % aproximado; sobre el Número de Personas en el Hogar los registros mostraron que la mayoría de los individuos vive solo (45 % aproximado), en agosto la mayoría de desocupados también vive solo (pero solo un 30 % aproximado), mientras que en noviembre la mayoría de individuos desocupado son en hogares con 3 personas (28 % aproximado); por último, la variable Edad mostró su mayoría de individuos en rango de 45 a 54 años (casi un 25 % en ambos meses), mientras que la mayoría de desocupados son de 25 a 34 años (casi 26 %) seguido de 15 a 24 años (24 % aproximado) en agosto y, en noviembre la mayoría de desocupados son de 15 a 24 años (30 % aproximado).

Ciertos resultados confirman lo expuesto por Astete B. [9], si bien sus resultados son obtenidos mediante la comparación de la tasa de desocupación en distintas variables, este trabajo presenta resultados en bruto de distintas variables según la Condición de desocupado, las conclusiones y resultados obtenidos son concordantes en las variables de trabajo. Una de las conclusiones principales de las Tablas de Contingencia son los mayores porcentajes de desocupado en Edad de 15 a 24 años, también se observó que el porcentaje de desocupado va disminuyendo conforme aumentan los años de los individuos, concordante con las mayores tasas de desocupación en los trimestres móviles de estudio. Con respecto a los Años de Estudio, se obtuvo que el menor porcentaje de desocupado son individuos sin estudio o hasta 8 años de estudio, y que el mayor porcentaje de desocupado son individuos de 9 a 12 años de estudio, resultados concordante y semejante con las tasas de desocupación por nivel educacional de la Región en los trimestres móviles de estudio, donde Astete B. expuso que la menor tasa de desocupación es en nivel primario respecto a los otros niveles y, la mayor tasa de desocupación es en nivel secundario y técnico.

En general, el análisis cruzado nos pone en frente cualidades específicas que presenta el individuo desocupado según su mayoría porcentual de las variables analizadas en los trimestres móviles en estudio y, al confirmar una relación de la Condición de Actividad con cada una de las variables en estudio (bajo análisis bivariado), se observó influencias de cada variable que se detallaron en ACS. Se observó que el individuo desocupado mayormente pertenece a la Provincia de Concepción, mayormente son hombres, mayormente son soltero(a), no proveedor, mayormente pertenecieron al Grupo 2 de Ocupación, mayormente individuos que su empleo perteneció a la actividad económica terciaria, principalmente tenían empleos dependientes, mayormente tienen rangos etarios jóvenes de 15 a 24 años o, tienen de 9 a 12 años de estudio. Algunas son cualidades esperadas o con una tendencia lógica, como la Edad, los Años de Estudio o la cualidad de Proveedor.

Los resultados de los Análisis de Correspondencia Simple (ACS), nos confirman una relación de cada una de las variables con la Condición de Actividad, pero en específico con que modalidad se corresponde más el individuo desocupado. Así, se logró encontrar los principales factores que están incidiendo en la Condición de desocupado. Los meses en estudio de la Región del Biobío mostraron resultados completamente equivalentes, no hubo variación en la relación que se encontró de las modalidades de cada una de las variables con el individuo desocupado en los meses centrales de agosto y de noviembre.

Se encontró relación de la Condición de desocupado con varias modalidades que ya tuvieron mayoría de porcentajes de desocupados en las Tablas de Contingencia, los resultados son vistos a modo bivariado, en específico se confirma una relación de la Condición de desocupado con Provincia de Concepción, con Estrato ciudades, con Sexo hombre (igual que la Condición de ocupado), con Estado Conyugal soltero(a), con no proveedor, con Empleo dependiente y, con individuos de 9 a 12 años de estudio.

Por otra parte, el ACS encontró relación con algunas modalidades distintas a la mayoría de porcentajes de desocupado en las Tablas de Contingencia de ciertas variables, en específico la Condición de desocupado se relacionó con Grupo 3 de Ocupación, con actividad económica secundaria, con hogares de 3 a 6 o más personas y, con rango etario de 15 a 24 años principalmente.

El Análisis de Correspondencia Múltiple (ACM) para la región, tiene una rápida conclusión, es que, al igual que en ACS, no hay grandes cambios de relaciones en ACM de mes de Agosto al mes de Noviembre. Las variaciones de las posiciones de las modalidades en las dimensiones analizadas son menores y casi nulas, otras ocurren en menor importancia por su menor inercia explicada ajustada.

Al no encontrar grandes cambios, se logra confirmar en cada análisis y en ambos meses una estructura de los factores, ya identificados en ACS, que inciden en la Condición de desocupado, en este caso de forma multivariada para la Región del Biobío. Se obtiene principalmente una constante proximidad con las modalidades de rango etario de 15 a 24 años y de 25 a 35 años, hogares de 3 a 6 o más personas, soltero(a) y no proveedor, resultando una relación global y de mayor importancia. En ambos meses aparece una relación al analizar la dimen-

sión 1 vs la dimensión 3, se aproxima a la Condición de desocupado con 15 a 24 años y de 25 a 34 años, de 3 a 6 o más personas en el hogar y soltero(a), la Situación de Empleo dependiente, aunque se ubica cerca del origen dada su alta frecuencia, representando una relación con el grupo próximo a la Condición de desocupado. Además, en ambos análisis se ubicaron en mismo cuadrante, aunque cercanas al origen por su alta frecuencia, la Provincia de Concepción y el Estrato ciudades, pero poseían sus más bajas medidas de discriminación lo que indica poca relación con las dimensiones.

Luego, en análisis de dimensiones de menos importancia (D2 vs D3) se mantienen próximos con desocupado los mismos rangos etarios, el número de personas en el hogar y el estado conyugal (a pesar que todas disminuyen su relación con los factores), ahora se ubican próximos y adquieren una relación más débil con el grupo mencionado, la Provincia del Biobío, Sexo hombre, el Grupo 2 de Ocupación y actividad secundaria para el mes de Agosto, mientras que para noviembre se alejan y quedan fuera del mismo cuadrante la Provincia del Biobío y actividad secundaria. Los individuos de 9 a 12 años de estudios también presentan una proximidad con este grupo en las mismas dimensiones, pero tienen su peor relación con los factores en ambos meses. El cambio de proximidad de la Provincia y la Rama de Actividad puede representar fenómenos puntuales de cada mes en estudio, un fenómeno estrictamente estacional o variaciones no medibles de la población.

El Análisis de Correspondencia Múltiple para cada Provincia tuvo relaciones similares que el ACM regional, sobre todo la Provincia de Concepción, las Provincias de Arauco y Biobío mantuvieron una base global con leves cambios en las relaciones o estructura principal. Con respecto a la comparación de un mes a otro, se presentaron resultados más dinámicos, sobre todo en la Provincia de Arauco, sin embargo la relación principal o global se observa claramente. En resumen, se logró apreciar una estructura similar de factores que inciden en la Condición de desocupado para cada Provincia, la que se discute a continuación.

La Provincia de Concepción mantiene una estructura de relación principal para la Condición de desocupado con rango etario de 15 a 24 años, hogares de 3 a 6 o más personas, soltero(a) o conviviente, y no proveedor. Automáticamente se aprecia un cambio de estructura que a nivel regional no se vio, el rango etario se acota solo de 15 a 24 años y el individuo conviviente se aproxima al grupo de relación principal (es la única modalidad que ingresa al cuadrante de interés) siendo una cualidad de la Provincia de Concepción. El análisis de la D1 vs D3, en ambos meses, aproxima y relaciona la Situación de Empleo dependiente al grupo antes mencionado, de forma idéntica que en análisis regional.

Un análisis de menor importancia (D2 vs D3) indica que, a la relación de Condición de desocupado con edad de 15 a 24 años, hogares de 3 a 6 o más personas, soltero(a) o conviviente, también se relaciona de manera más débil el individuo hombre, el Grupo 2 de Ocupación y actividad secundaria en ambos meses; de 9 a 12 años de estudio solo en agosto. Los resultados de Concepción son casi equivalentes al resultado regional, existe una relación más débil con el mismo Sexo, Grupo de Ocupación y Rama de Actividad. Además, para Concepción se confirma de mejor forma la relación con individuos de 9 a 12 años de estudio, dado que la variable involucrada posee su mejor representación en dichas dimensiones.

La Provincia de Arauco presentó una estructura de relación principal o global de la Condición de desocupado con rango etario de 15 a 24 años y de 25 a 34 años, hogares de 4 a 6 o más personas (inclusive mejor representado de 6 o más personas en noviembre), soltero(a) (conviviente solo en agosto), no proveedor e individuos de 9 a 12 años de estudio. En Arauco se destaca que llega como parte de una estructura principal la relación de individuos de 9 a 12 años de estudio con desocupado y el grupo próximo, en general se mantiene en las 3 dimensiones analizadas; además la relación principal tiene más cambios entre meses, por ejemplo en Estado Conyugal y Número de Personas en el Hogar, que llama mucho la atención ya que se mantiene pero en hogares con mayor número de personas, en relación a la Región del Biobío y con Provincia de Concepción. El análisis de la D1 vs D3, reitera la relación de desocupado con el grupo próximo e incluye la Situación de Empleo dependiente, solo para mes de agosto.

El análisis de menor importancia (D2 vs D3) mantiene la relación principal de desocupado con el grupo próximo, y en agosto solo se relaciona con individuos hombre. En cambio, en mes de noviembre desocupado y el grupo próximo se relaciona con Estrato rural y Grupo 2 de Ocupación. Se destaca que actividad primaria estuvo próxima en agosto, pero con baja relación con los factores y, en noviembre estuvo muy cerca de estar relacionada con el grupo principal. El análisis realizado en Arauco es el que más difiere de un mes a otro (o de una estación a otra), dejando en evidencia el mayor dinamismo de su población, inclusive varía en muchas medidas con respecto a Concepción y la Región.

La Provincia del Biobío presentó una estructura de relación principal de la Condición de desocupado con rango etario de 15 a 24 años, hogares de 3 a 6 o más personas, soltero(a) (conviviente solo en agosto) y no proveedor. Rápidamente se concluye que Biobío es más parecida a la Provincia de Concepción, mantiene mismo rango etario y Número de Personas en el Hogar. En análisis de D1 vs D3, próximo a desocupado y al grupo relacionado, se agrega la relación con Empleo dependiente, como ha sido común en la Región y Provincia de Concepción. Además, bajo su mejor relación con los factores, de 13 a 18 años de estudio aparece próxima al grupo y también ciudades pero más cerca del origen (dada su alta frecuencia) en mes de agosto. Para mes de noviembre, los años de estudio se mantiene en el borde de ingresar al cuadrante que mantiene al grupo de relación principal. Esto puede explicar un mayor movimiento de individuos estudiantes de educación superior o egresados, que en época habitual de clases no logran trabajar, mientras que en época de fin de año se abren más puestos de trabajo part-time (o tiempo completo) que buscan mayormente este grupo de individuos.

El análisis de menor importancia (D2 vs D3) sigue la tendencia de mantener en gran parte la relación principal, y tiene un parecido al análisis regional. Se relaciona de forma más débil al grupo principal, RAU, individuo hombre, el Grupo 2 de Ocupación y actividad secundaria (solo en agosto), teniendo gran parecido al análisis regional y de Concepción. Además, de 9 a 12 años de estudio es próximo a desocupado y al grupo en ambos meses, con su menor relación con las dimensiones, equivalente a análisis regional.

Los resultados de Astete B. [9] ya advertían sobre la mayor tasa de desocupado en individuos de 15 a 24 años en la Región y en especial para cada Provincia, en este trabajo se concluyó y encontró una relación principal o una estructura para la Condición de desocupado en las 3 Provincias que incluye dicho rango etario. Además, aprovechando la mayor información posible de la encuesta, se intentó abarcar una mayor cantidad de variables sociodemográficas e incluir variables económicas, con las que se obtuvieron nuevos resultados en forma bivariada y multivariada, logrando identificar nuevos factores y analizando una estructura de factores que inciden en el desempleo (de la Región y sus Provincias). Lo novedoso fue la inclusión de las variables económicas, las que muestran una incidencia en el desempleo y una estructura conjunta, en base a la información de su última actividad en la ocupación.

Queda en evidencia una estructura de los factores que inciden en el desempleo para la Región y para cada Provincia, observando algunas variaciones en éstas. Bajo este trabajo, es posible analizar de forma más específica o tomar más conocimiento de esta estructura presente en la Región y Provincias, dado que son factores medibles obtenidos de la Encuesta Nacional de Empleo (ENE) y que tuvieron relación con la Condición de desocupado.

Se destaca que es posible realizar este análisis en trimestres más actuales, para analizar si mantiene la misma estructura de factores que inciden en el desempleo o si hay cambios, es ideal que los estudios posteriores se realicen en condiciones sociales normales. Se recomienda tener en cuenta que, el trabajo con 2 meses consecutivos involucra información repetida, ya que los resultados mensuales provienen de trimestres móviles.

Se podría realizar estudios longitudinales en que se tome especial atención a los meses centrales de Febrero, Mayo, Agosto y Noviembre, dado que representan mayormente los cambios de estaciones del año; incluir nuevas variables sociodemográficas o económicas como la cantidad de ofertas de empleo y la demanda de empleo. Otros análisis, enfocados a la predicción de la Condición de Actividad de los individuos, realizar modelos logísticos o aplicar algoritmos de clasificación como árboles de clasificación.

Capítulo 7

Anexo.

7.1. ACM Provincia de Concepción.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,0094	0,1386	0,0295
Personas en el Hogar	0,6722	0,0999	0,2716
Edad	0,5702	0,1608	0,0885
Sexo	0,0846	0,0720	0,4308
Años de Estudio	0,2184	0,3942	0,1259
Estado Conyugal	0,4759	0,0542	0,0720
Proveedor	0,5931	0,0464	0,0745
Grupo de Ocupación	0,0725	0,5064	0,0633
Rama de Actividad	0,0947	0,2267	0,2664
CISE	0,0611	0,0561	0,1381
Condición Actividad	0,0628	0,1036	0,1449
Promedio (Inercia)	0,2650	0,1690	0,1550

Tabla 7.1: Medidas de discriminación, ACM para Provincia de Concepción, agosto.

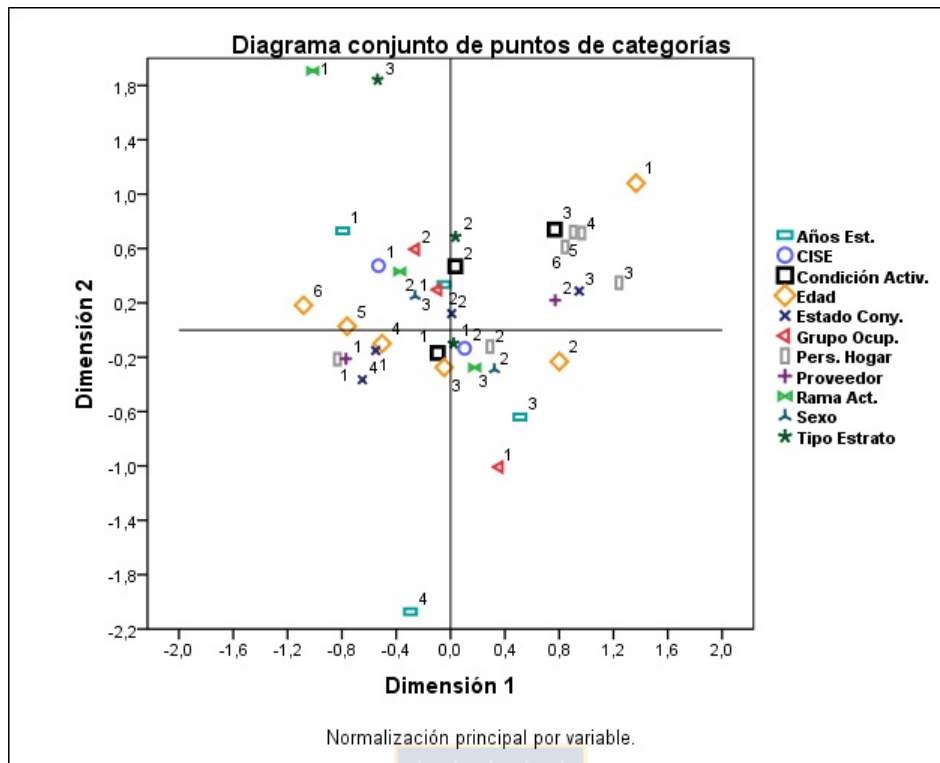


Figura 7.1: ACM para Provincia de Concepción, D1 vs D2 agosto.

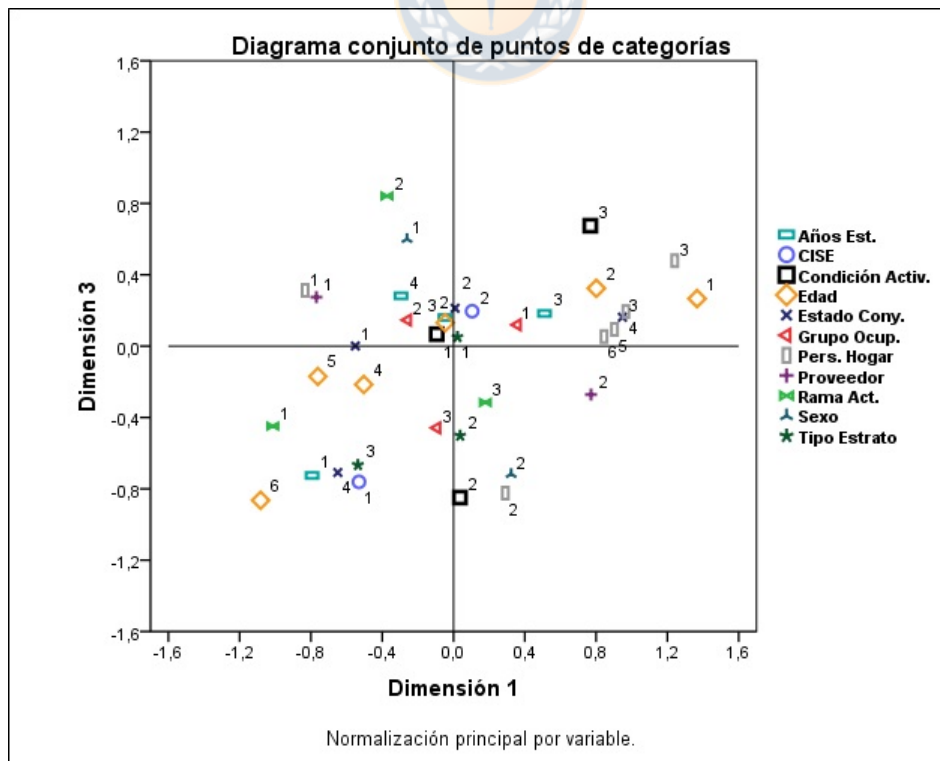


Figura 7.2: ACM para Provincia de Concepción, D1 vs D3 agosto.

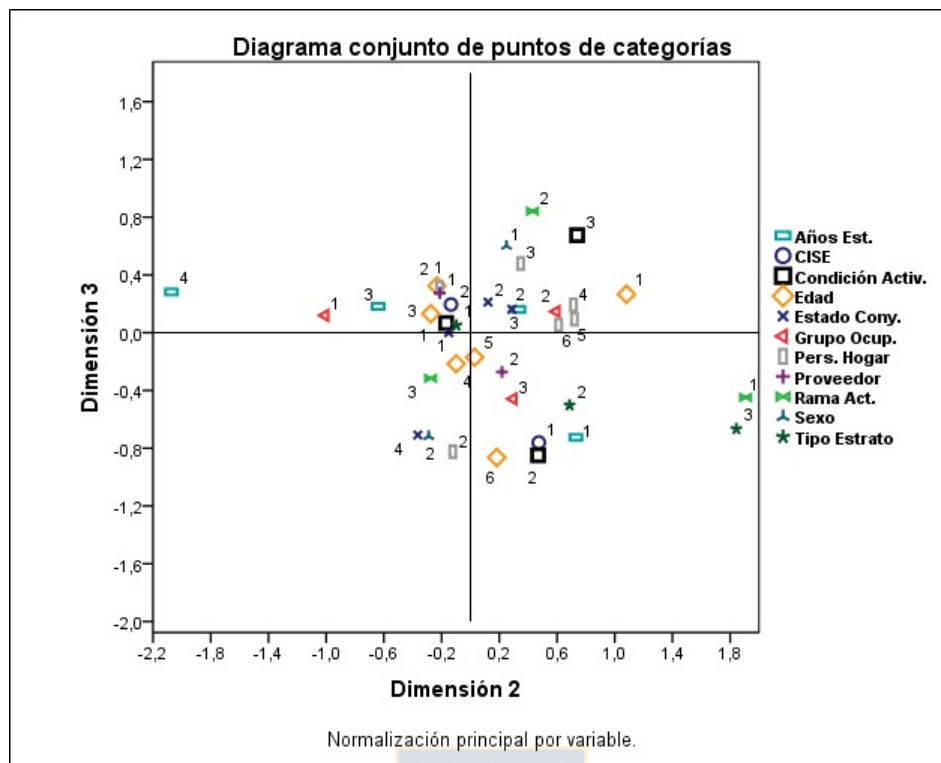


Figura 7.3: ACM para Provincia de Concepción, D2 vs D3 agosto.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,0123	0,0751	0,0449
Personas en el Hogar	0,6642	0,1303	0,3019
Edad	0,5832	0,1538	0,0748
Sexo	0,0829	0,2116	0,3104
Años de Estudio	0,2153	0,2791	0,2441
Estado Conyugal	0,4721	0,0735	0,0355
Proveedor	0,5711	0,0189	0,1185
Grupo de Ocupación	0,0955	0,4306	0,2167
Rama de Actividad	0,0760	0,3184	0,1653
CISE	0,0683	0,0173	0,0801
Condición Actividad	0,0985	0,1185	0,0739
Promedio (Inercia)	0,2672	0,1661	0,1515

Tabla 7.2: Medidas de discriminación, ACM para Provincia de Concepción, noviembre.

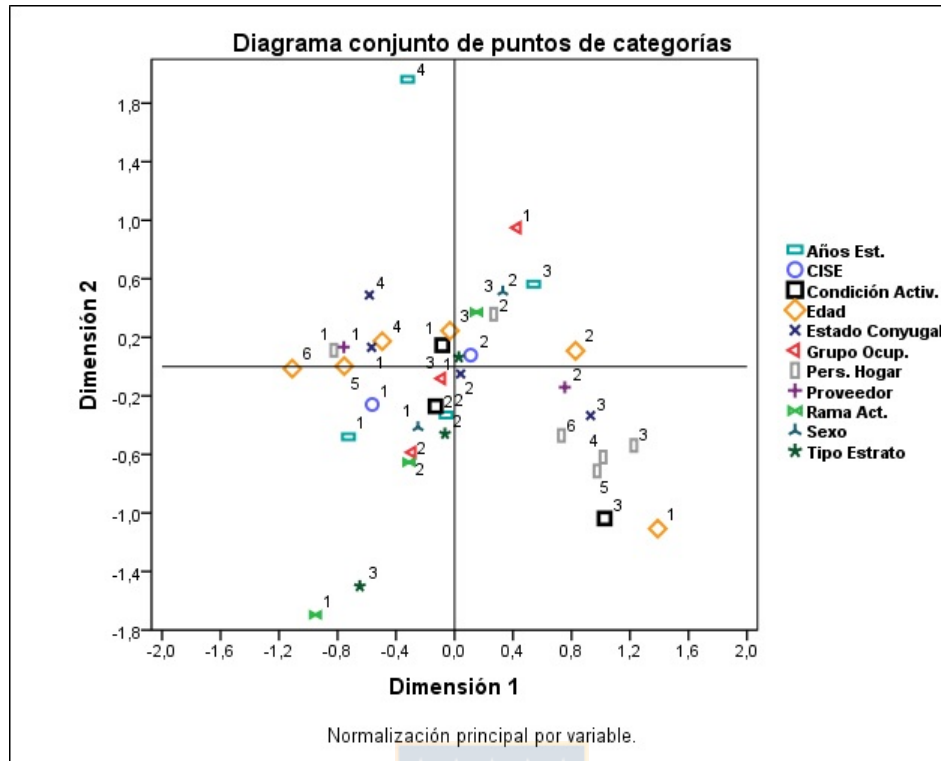


Figura 7.4: ACM para Provincia de Concepción, D1 vs D2 noviembre.

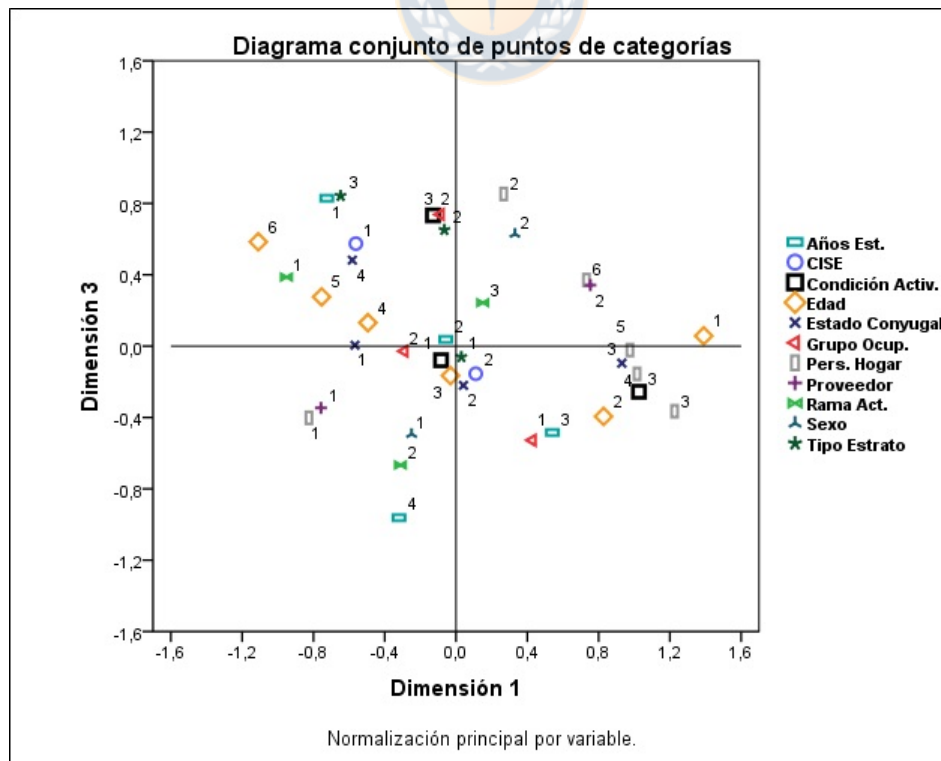


Figura 7.5: ACM para Provincia de Concepción, D1 vs D3 noviembre.

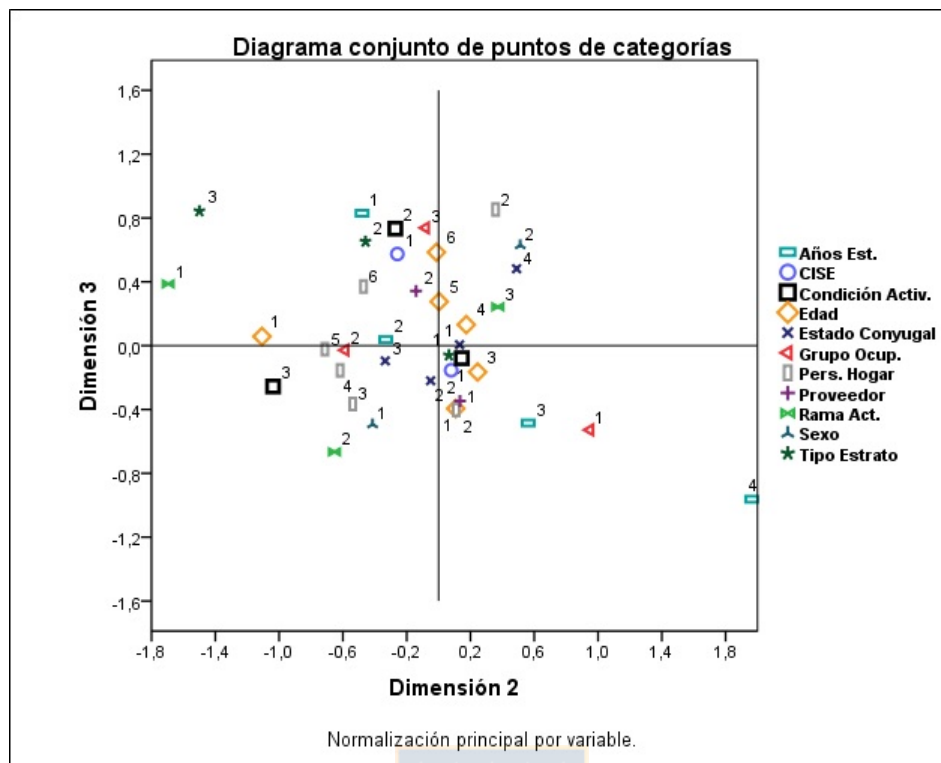


Figura 7.6: ACM para Provincia de Concepción, D2 vs D3 noviembre.

7.2. ACM Provincia de Arauco.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,1438	0,1779	0,0111
Personas en el Hogar	0,5241	0,2694	0,2971
Edad	0,4203	0,3508	0,1927
Sexo	0,1705	0,1045	0,2605
Años de Estudio	0,4179	0,1260	0,0717
Estado Conyugal	0,2729	0,1956	0,0965
Proveedor	0,4421	0,0963	0,1742
Grupo de Ocupación	0,2242	0,2674	0,0511
Rama de Actividad	0,2744	0,2507	0,0352
CISE	0,1541	0,0400	0,3697
Condición Actividad	0,0328	0,0609	0,2027
Promedio (Inercia)	0,2797	0,1763	0,1602

Tabla 7.3: Medidas de discriminación, ACM para Provincia de Arauco, agosto.

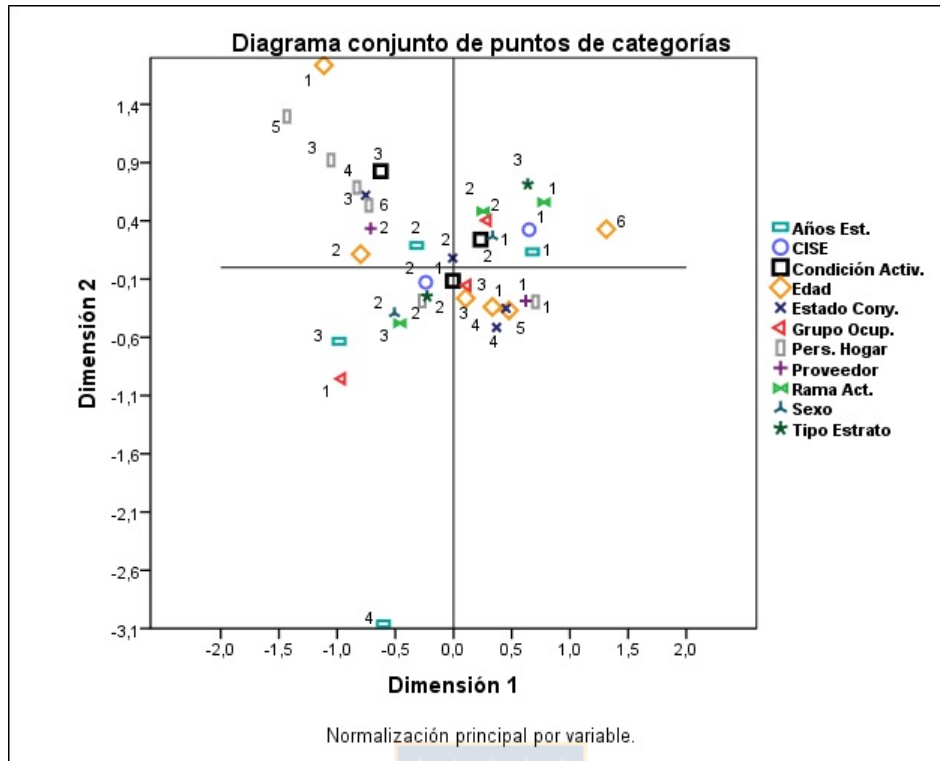


Figura 7.7: ACM para Provincia de Arauco, D1 vs D2 agosto.

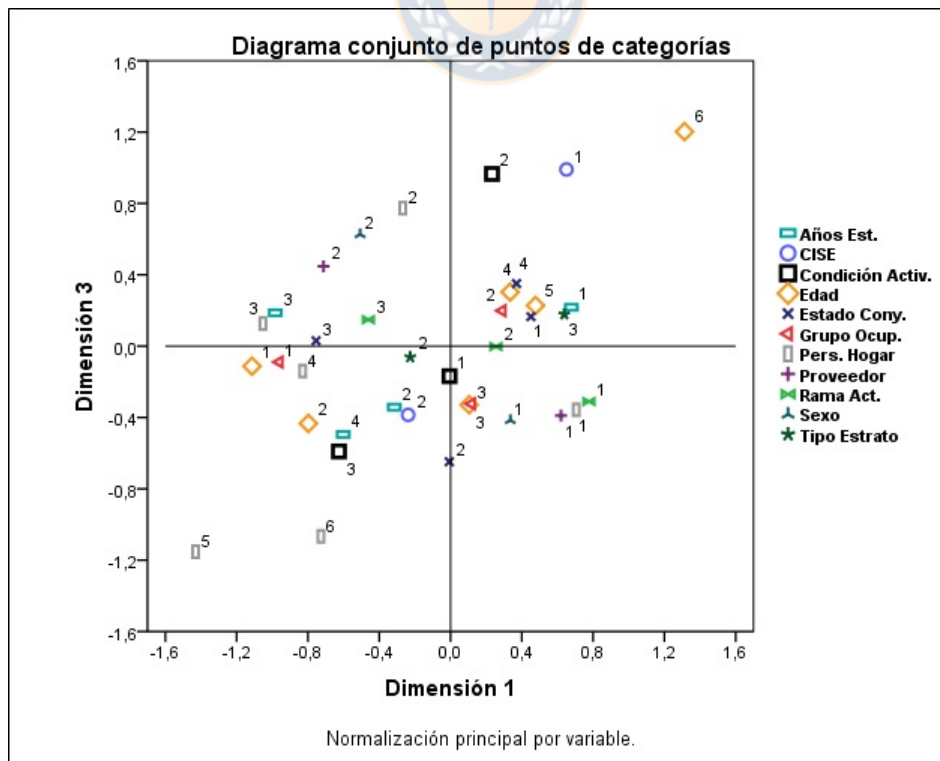


Figura 7.8: ACM para Provincia de Arauco, D1 vs D3 agosto.

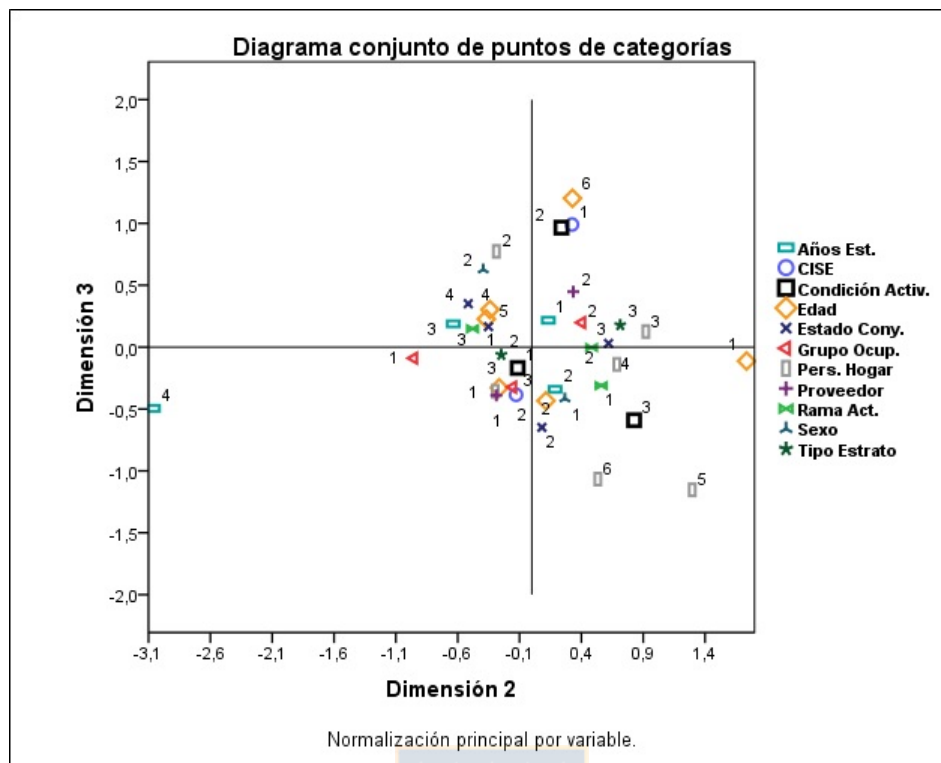


Figura 7.9: ACM para Provincia de Arauco, D2 vs D3 agosto.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,0792	0,1655	0,0405
Personas en el Hogar	0,5466	0,2699	0,4688
Edad	0,4516	0,3297	0,0943
Sexo	0,1791	0,2335	0,2208
Años de Estudio	0,4011	0,1525	0,1424
Estado Conyugal	0,3517	0,1954	0,0283
Proveedor	0,4648	0,0263	0,2737
Grupo de Ocupación	0,2121	0,2554	0,1568
Rama de Actividad	0,2142	0,3392	0,0149
CISE	0,1136	0,0012	0,1584
Condición Actividad	0,0457	0,0514	0,0668
Promedio (Inercia)	0,2782	0,1836	0,1514

Tabla 7.4: Medidas de discriminación, ACM para Provincia de Arauco, noviembre.

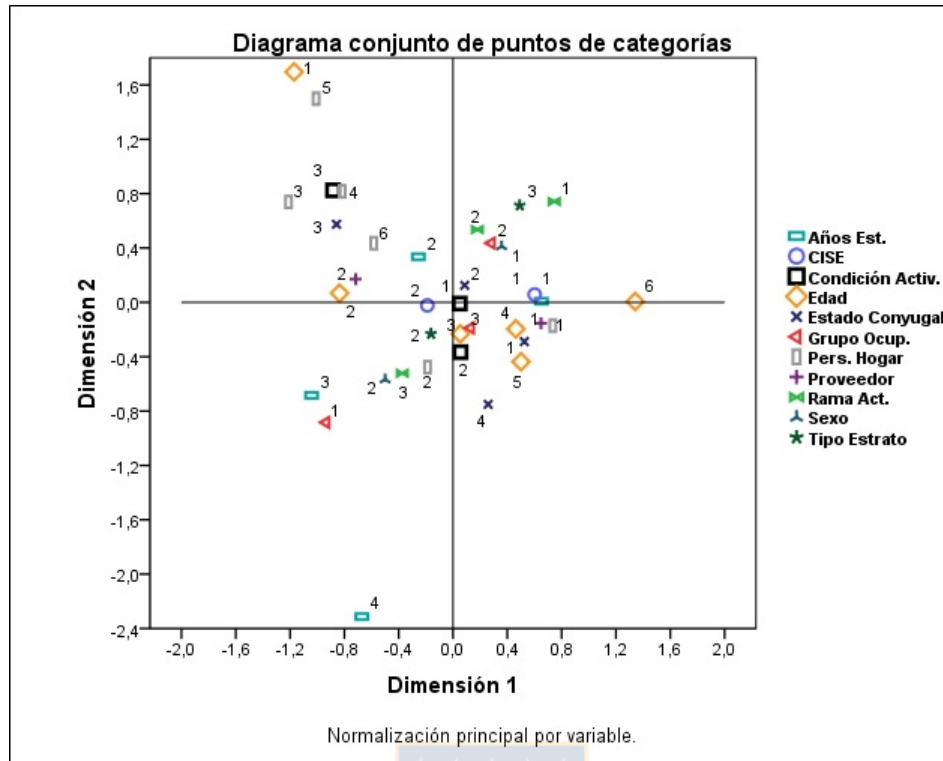


Figura 7.10: ACM para Provincia de Arauco, D1 vs D2 noviembre.

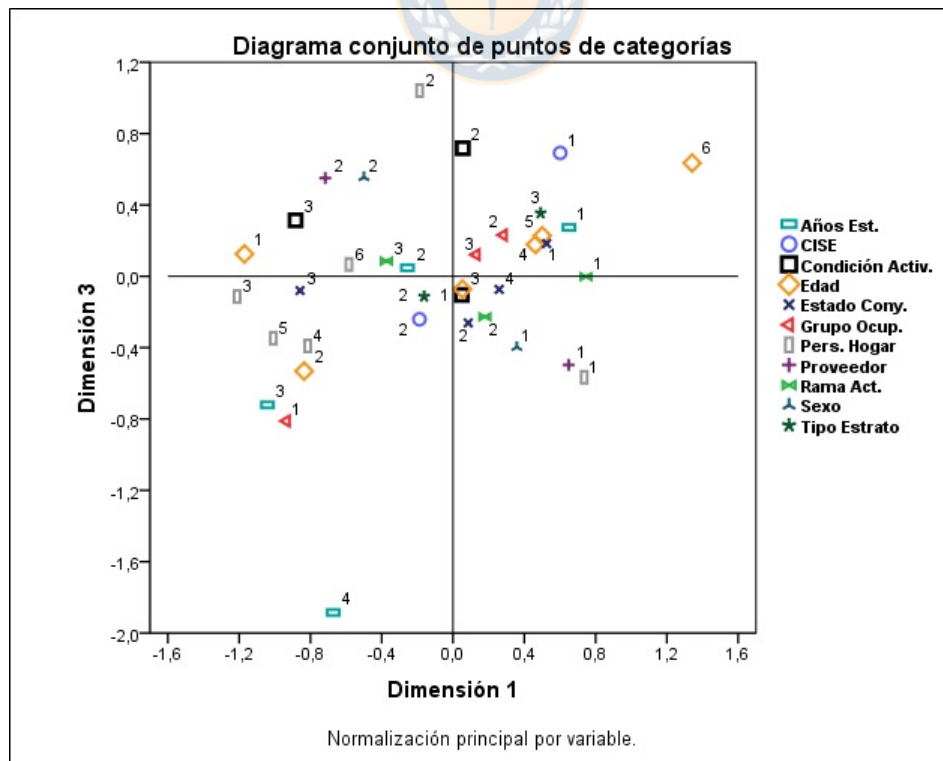


Figura 7.11: ACM para Provincia de Arauco, D1 vs D3 noviembre.

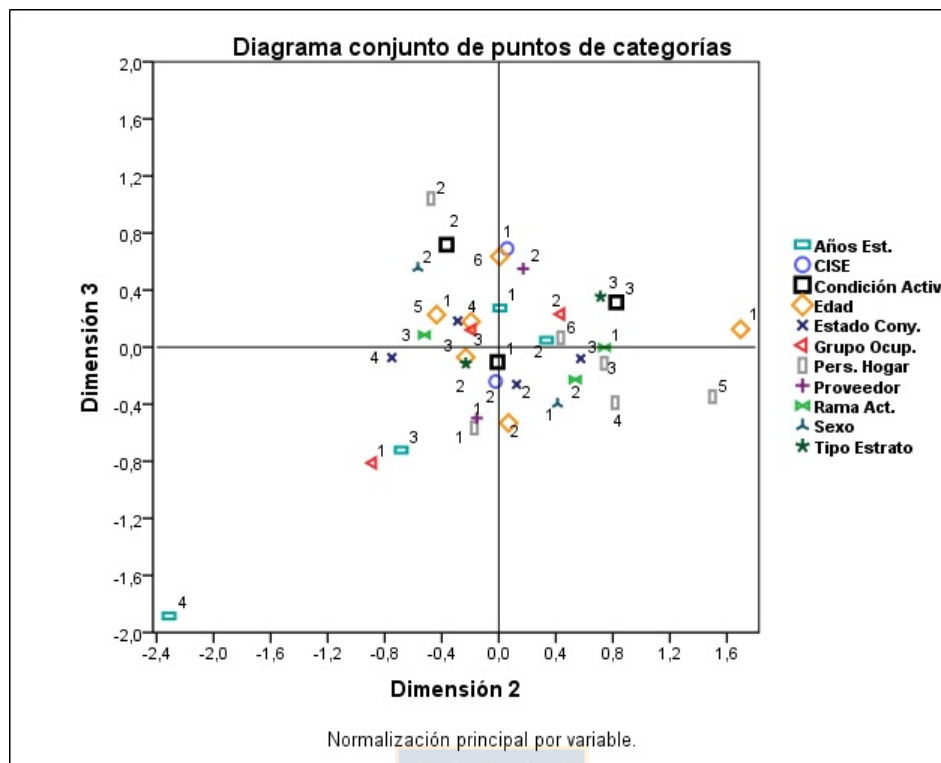


Figura 7.12: ACM para Provincia de Arauco, D2 vs D3 noviembre.

7.3. ACM Provincia del Biobío.

VARIABLES	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,0947	0,1202	0,1864
Personas en el Hogar	0,5831	0,3676	0,2241
Edad	0,4357	0,3413	0,0820
Sexo	0,2145	0,1551	0,2081
Años de Estudio	0,3224	0,1599	0,2794
Estado Conyugal	0,3515	0,1916	0,0225
Proveedor	0,5177	0,0979	0,1104
Grupo de Ocupación	0,2124	0,2566	0,1609
Rama de Actividad	0,2222	0,2165	0,3264
CISE	0,1006	0,0016	0,0466
CondiciónActividad	0,0103	0,0489	0,0042
Promedio (Inercia)	0,2787	0,1779	0,1501

Tabla 7.5: Medidas de discriminación, ACM para Provincia de Biobío, agosto.

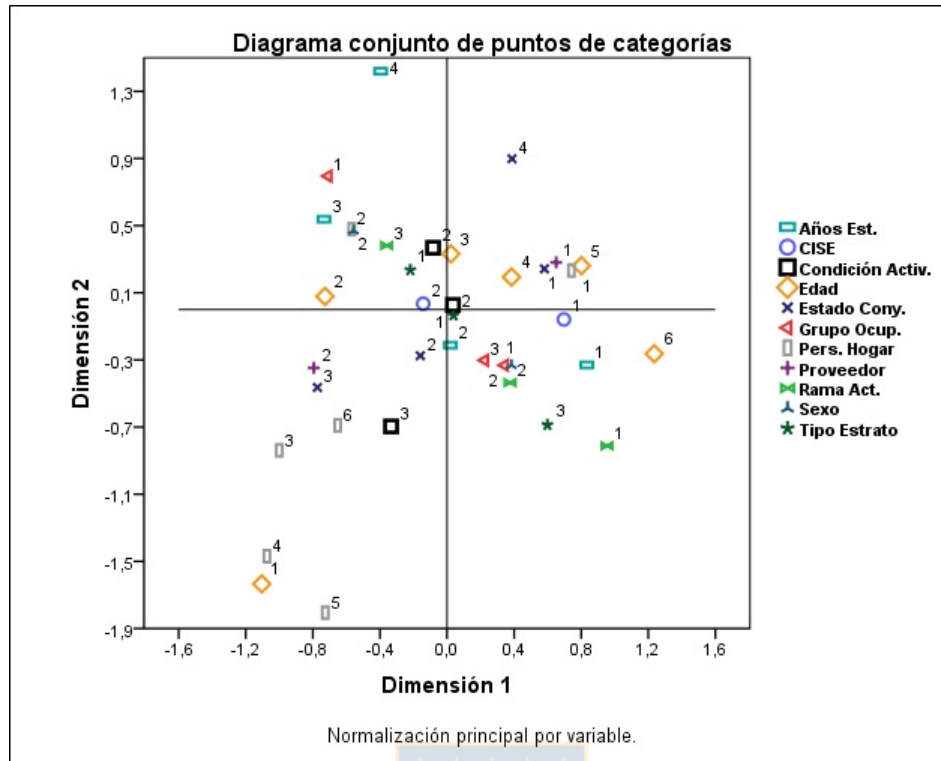


Figura 7.13: ACM para Provincia del Biobío, D1 vs D2 agosto.

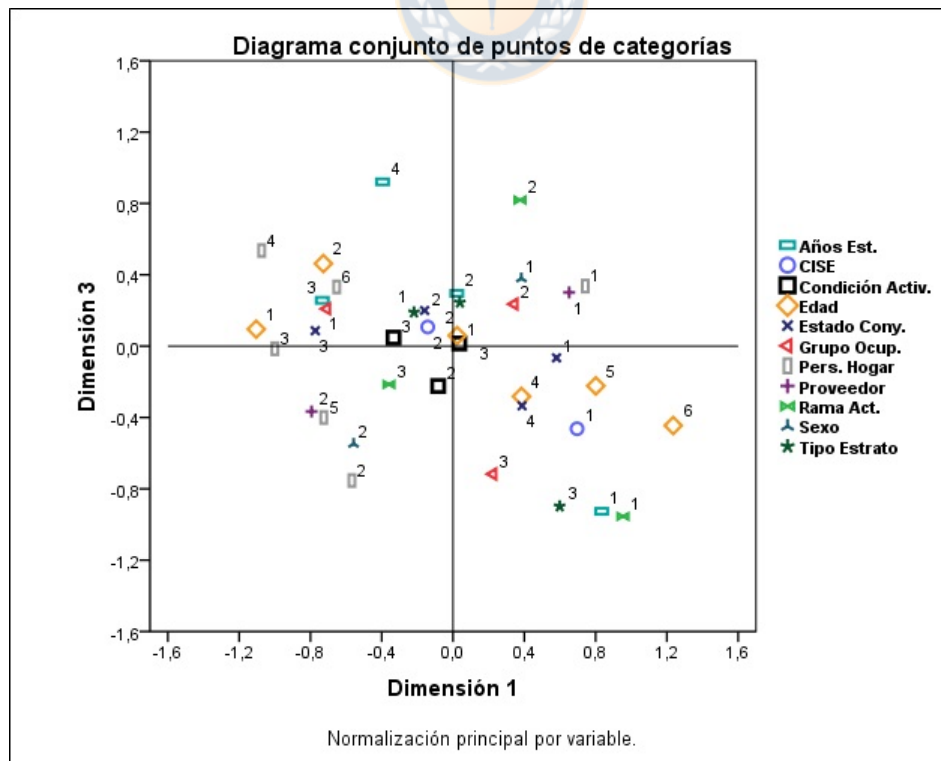


Figura 7.14: ACM para Provincia del Biobío, D1 vs D3 agosto.

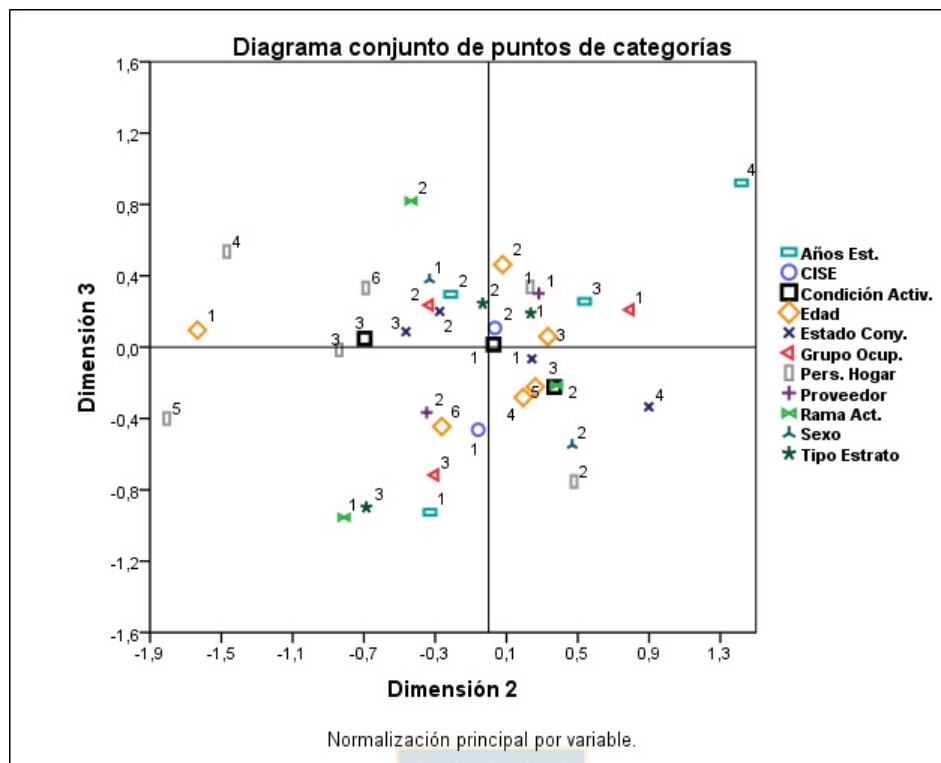


Figura 7.15: ACM para Provincia del Biobío, D2 vs D3 agosto.

Variables	Dimensión 1	Dimensión 2	Dimensión 3
Tipo Estrato	0,0948	0,2580	0,1647
Personas en el Hogar	0,5920	0,2732	0,1905
Edad	0,4695	0,3144	0,1366
Sexo	0,2372	0,0346	0,2445
Años de Estudio	0,3432	0,2238	0,2461
Estado Conyugal	0,3603	0,1445	0,0181
Proveedor	0,5396	0,1146	0,0199
Grupo de Ocupación	0,1913	0,3409	0,2845
Rama de Actividad	0,2091	0,2736	0,3742
CISE	0,0852	0,0028	0,0067
Condición Actividad	0,0140	0,0415	0,0098
Promedio (Inercia)	0,2851	0,1838	0,1541

Tabla 7.6: Medidas de discriminación, ACM para Provincia de Biobío, noviembre.

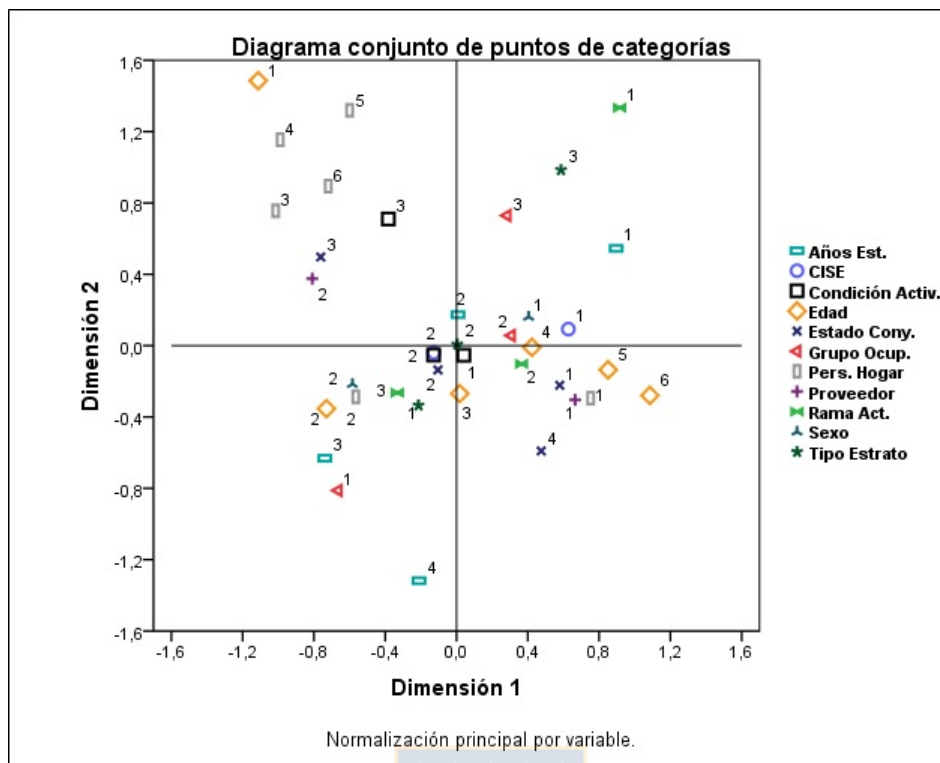


Figura 7.16: ACM para Provincia del Biobío, D1 vs D2 noviembre.

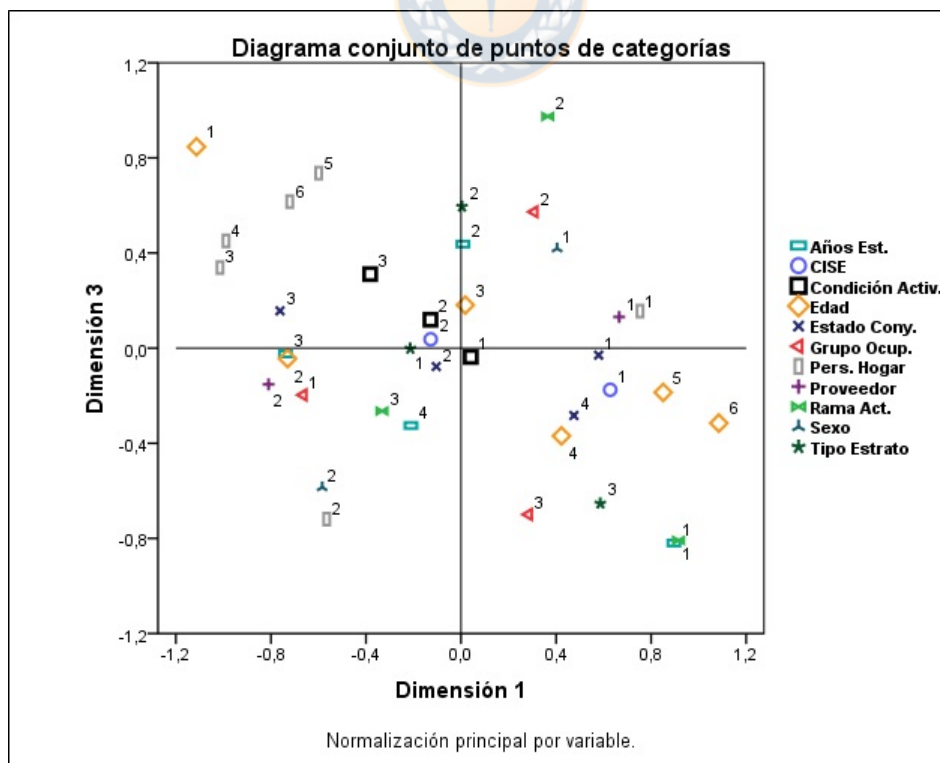


Figura 7.17: ACM para Provincia del Biobío, D1 vs D3 noviembre.

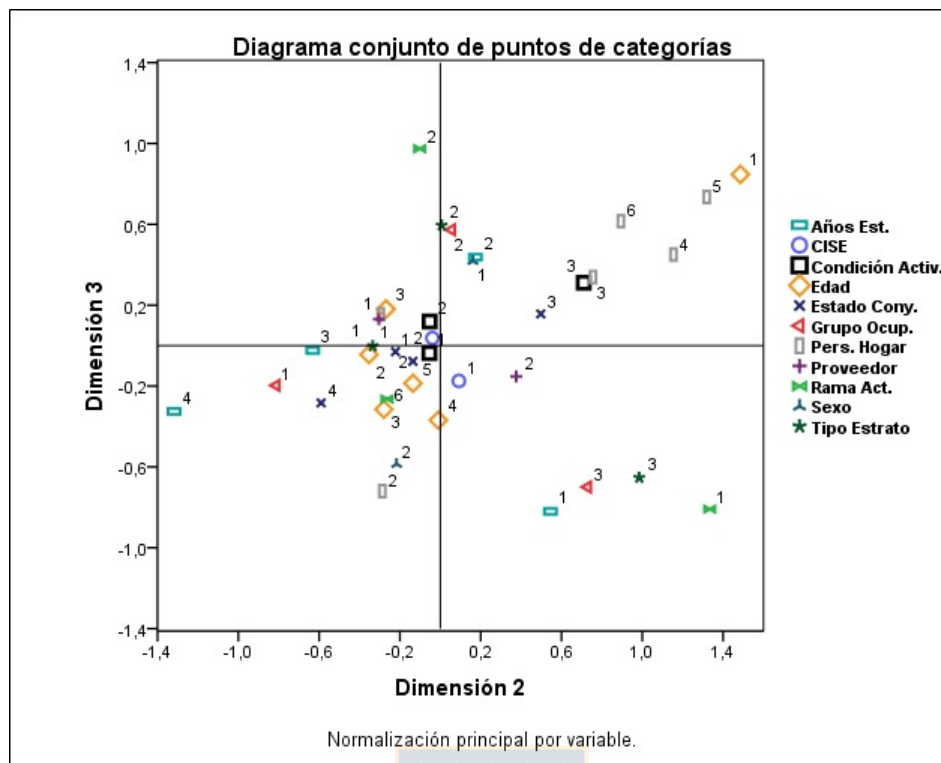


Figura 7.18: ACM para Provincia del Biobío, D2 vs D3 noviembre.

Bibliografía

- [1] <https://bancodatosene.ine.cl/>
- [2] <https://www.gestiopolis.com/que-es-desempleo-que-lo-causa-que-tipos-hay-como-se-mide/>
- [3] http://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-04622016000200049
- [4] <http://www.ine.cl/docs/default-source/laborales/ene/antecedentes-metodologicos/enfoque-estad%C3%A9stico-lanzamiento-de-la-nueva-encuesta-nacional-de-empleo.pdf?sfvrsn=4>
- [5] <http://www.ine.cl/docs/default-source/laborales/ene/antecedentes-metodologicos/separata-t%C3%A9cnica-nueva-encuesta-nacional-de-empleo.pdf?sfvrsn=4>
- [6] http://historico.ine.cl/canales/chile_estadistico/mercado_del_trabajo/empleo/metodologia/pdf/031110/manual_metodologico031110.pdf
- [7] http://historico.ine.cl/canales/chile_estadistico/mercado_del_trabajo/empleo/metodologia/pdf/glosarioNENE.pdf
- [8] [https://regiones.ine.cl/biobio/estadisticas#Boletines_Informativos/Encuesta_nacional_de_empleo_-_ENE/Encuesta_Nacional_Empleo_ENE_\(Tabulados\)](https://regiones.ine.cl/biobio/estadisticas#Boletines_Informativos/Encuesta_nacional_de_empleo_-_ENE/Encuesta_Nacional_Empleo_ENE_(Tabulados))
- [9] **Astete G. Bárbara.** (2017). Análisis de factores que inciden en la tasa de desocupación en la región del Bío Bío. Universidad de Concepción.
- [10] **Carter C. David.** (2015). Aplicación de algunas técnicas de análisis multivariado de la encuesta de demanda turística de los establecimientos de alojamiento turístico de la región del Bío Bío en el periodo estival 2015. Universidad de Concepción.
- [11] **Riquelme A. Gloria.** (2017). Encuesta sobre sexualidad en adolescentes de liceos de zonas urbano-rural de la octava región mediante análisis de correspondencias. Universidad de Concepción.
- [12] **Aguilera del Pino, Ana María.** (2005). Análisis de tablas de contingencia bidimensionales. Recuperado de web: <https://www.ugr.es/focana/dclasif/aaguilera.pdf> [04-2019]
- [13] **Salvador Figueras, M** (2003): “Análisis de Correspondencias” [en línea] 5campus.com, Estadística <http://www.5campus.com/leccion/correspondencias> [06-02-2019]

- [14] **Escofier Brigitte, Pagès Jérôme.** (1992). Análisis factoriales simples y múltiples: objetivos, métodos e interpretación. Edición española Bilbao, España. Edit. UPV.
- [15] **Gallardo San Salvador, José Ángel.** Análisis de Correspondencias Simple. Universidad de Granada. Recuperado de web: <https://www.ugr.es/gallardo/pdf/acs-g.pdf> [28-05-2019]
- [16] **Colmenares Lacruz, Gerardo.** (2012). Análisis de Correspondencia. Universidad de los Andes. Recuperado de web: http://webdelprofesor.ula.ve/economia/gcolmen/programa/economia/analisis_correspondencia.pdf [22-02-2019]
- [17] **Peña, Daniel.** (2002). Análisis de datos multivariantes. McGraw-Hill Interamericana de España, S A U. Edit. Concepción Fernández Madrid.
- [18] **Díaz Monroy, Luis Guillermo.** (2007). Estadística multivariada: Inferencia y métodos. Bogotá, Colombia. Editorial: Universidad Nacional de Colombia, Facultad de Ciencias. Segunda edición.
- [19] **Joaristi, Luis.** (1999). Análisis de correspondencias. Madrid, España. Editorial: LA MURALLA S.A.
- [20] **Cuadras, Carles M.** (2014). Nuevos métodos de análisis multivariantes. Barcelona, España. Editorial: CMC Editions.
- [21] **Greenacre, Michael.** (2008). La práctica del análisis de correspondencias. Edición en español, Fundación BBVA.
- [22] **De la Fuentes Fernández, Santiago.** (2011). Análisis correspondencias simples y múltiples. Madrid: Universidad Autónoma de Madrid.
- [23] **Vivanco, Manuel.** (1999). Análisis estadístico multivariable. Teoría y práctica. Santiago, Chile. Editorial Universitaria S.A. Primera edición.
- [24] **Fernández Gómez, María José.** (2005). Análisis Factorial de Correspondencias. Universidad de Salamanca. Programa de Doctorado: Estadística Multivariante Aplicada. España.