



## **Anexo 4: Productos de la consultoría**

### **Anexo 4.2: Manual para el manejo del System Analysis Statistical v9.2 (SAS) para análisis avanzado de datos**

Fomento a la construcción de capacidades en relación a las políticas públicas apoyadas  
por Contratos de Reforma Sectorial (CRS) en Bolivia (DCI-ALA/2014/025-058)

**Fortalecimiento de la capacidad institucional en los sectores de  
desarrollo integral con coca, tráfico ilícito de drogas y seguridad  
alimentaria para una eficiente gestión del apoyo presupuestario  
sectorial en Bolivia (DITISA)**

***Manual para el manejo del System Analysis Statistical  
v9.2 (SAS) para análisis avanzado de datos  
Informe final de misión de Corta Duración***

Contrato n° DCI/LA/2017/392-699  
N° de identificación: EuropeAid/138320/IH/SER/BO



Proyecto Financiado por la  
Unión Europea



Implementada por el consorcio:

AGRER — AECOM — TRANSTEC

*La Paz / Bruselas, enero 2019*

---

**Disclaimer:**

Este informe ha sido elaborado por el consorcio AGRER/AECOM/TRANSTEC con financiamiento de la Unión Europea. Las opiniones aquí expresadas son del consultor y no expresan necesariamente las de la Comisión Europea.

---



## AGRADECIMIENTOS

Ministerio de Desarrollo Rural y Tierras

[Dirección General de Planificación y Observatorio Agropecuario Productivo]

Programa Financiado por Unión Europea.

Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial.

DITISA-UE. LA/2017/392-699



## CONTENIDO

|                                                                                             |           |
|---------------------------------------------------------------------------------------------|-----------|
| <b>Lista de las figuras .....</b>                                                           | <b>6</b>  |
| <b>Lista de las tablas .....</b>                                                            | <b>6</b>  |
| <b>1. Organización de datos .....</b>                                                       | <b>1</b>  |
| 1.1.Ventanas del SAS.....                                                                   | 1         |
| 1.2.Introducción de datos en ventana Editor .....                                           | 1         |
| 1.3.Ejecución de los procedimientos y sentencias.....                                       | 3         |
| <b>2. Análisis univariable .....</b>                                                        | <b>4</b>  |
| 2.1.Procedimiento UNIVARIATE.....                                                           | 4         |
| 2.2.Procedimiento MEANS.....                                                                | 10        |
| 2.3.Procedimiento FREQ .....                                                                | 11        |
| <b>3. Análisis bivariable .....</b>                                                         | <b>14</b> |
| 3.1.Entre dos variables cualitativas: tablas de contingencia y prueba de chi cuadrado. .... | 14        |
| 3.2.Para una variable cuantitativa y otra cualitativa: análisis de varianza. ....           | 16        |
| Análisis de varianza. ....                                                                  | 16        |
| Procedimientos para el análisis de varianza en SAS .....                                    | 16        |
| 3.3.Entre dos variables cuantitativas: correlación y regresión lineal simple. ....          | 20        |
| Análisis de correlación .....                                                               | 20        |
| Regresión lineal.....                                                                       | 23        |
| <b>4. Análisis multivariable .....</b>                                                      | <b>25</b> |
| 4.1.Análisis de regresión múltiple .....                                                    | 25        |
| Supuestos del análisis de regresión .....                                                   | 28        |
| 4.2.Análisis de correspondencia .....                                                       | 31        |
| Análisis de correspondencia múltiple .....                                                  | 34        |
| 4.3.Análisis de agrupación .....                                                            | 36        |
| Cluster Método Wards (Variables cuantitativas) .....                                        | 36        |
| Componentes principales .....                                                               | 39        |
| Cluster Método Average (Variables cualitativas).....                                        | 43        |
| 4.4.Análisis de regresión logística .....                                                   | 48        |
| <b>Bibliografía .....</b>                                                                   | <b>52</b> |



## LISTA DE LAS FIGURAS

|                                              |   |
|----------------------------------------------|---|
| Figura 1. Vista de las ventanas de SAS ..... | 1 |
|----------------------------------------------|---|

## LISTA DE LAS TABLAS

|                                                                                                                                                                                                                |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabla 1. Detalles descritos sobre procedimientos y sentencias de DATA y PRINT .....                                                                                                                            | 2  |
| Tabla 2. Procedimiento DATA e información de la Encuesta de papa en Cochabamba 2017 .....                                                                                                                      | 5  |
| Tabla 3. Procedimientos y sentencias de UNIVARIATE para la variable rend1718 (Rendimiento de papa campaña 2017-2018 (t/ha)) (OAP-MDRyT, 2018) .....                                                            | 5  |
| Tabla 4. Salidas del procedimiento y sentencia de Univariate de la Tabla 3. ....                                                                                                                               | 5  |
| Tabla 5. Procedimientos y sentencias de UNIVARIATE para el rendimiento de papa (t/ha) .....                                                                                                                    | 7  |
| Tabla 6. Salidas del procedimiento y sentencia de Univariate de la Tabla 5 (solo del rendimiento y las distribuciones. ....                                                                                    | 8  |
| Tabla 7. Procedimientos y sentencias de MEANS para la variable rendimiento de papa (t/ha) .....                                                                                                                | 10 |
| Tabla 8. Salidas del procedimiento y sentencia de MEANS de la Tabla 7. ....                                                                                                                                    | 10 |
| Tabla 9. Procedimientos y sentencias de FREQ para la variable P23 ¿Utilizo algún agroquímico en la campaña 2016-2017? .....                                                                                    | 11 |
| Tabla 10. Salidas del procedimiento y sentencia de FREQ de la Tabla 9. ....                                                                                                                                    | 12 |
| Tabla 11. Procedimientos y sentencias de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? .....                                                                                 | 13 |
| Tabla 12. Salidas del procedimiento y sentencia de FREQ de la Tabla 11. ....                                                                                                                                   | 13 |
| Tabla 13. Procedimientos y sentencias de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? por variedad cultivada .....                                                          | 14 |
| Tabla 14. Salidas del procedimiento y sentencia de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? por variedad cultivada Tabla 13. ....                                       | 15 |
| Tabla 15. Al mismo tiempo se debe diferenciar las características del análisis de varianza en el muestreo y diseño experimental (Leon y Quiroz, 1994): .....                                                   | 16 |
| Tabla 16. Procedimientos y sentencias de GLM para la variable rend1718 (Rendimiento de papa t/ha) y variedad cultivada.....                                                                                    | 17 |
| Tabla 17. Salidas del procedimiento y sentencia de GLM para la variable rend1718 (Rendimiento de papa t/ha) y variedad cultivada.....                                                                          | 17 |
| Tabla 18. Procedimiento y sentencia de CORR (análisis de correlación) para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1..... | 21 |
| Tabla 19. Salida del procedimiento y sentencia de CORR para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1. ....               | 21 |
| Tabla 20. Procedimiento y sentencia de SORT y CORR para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1 por variedad. ....      | 22 |
| Tabla 21. Procedimiento y sentencia de REG para la variable rend1718 (Rendimiento de papa t/ha) y superficie cultivada. Y gráfica de dispersión de valores observados y estimados sobre la superficie. ....    | 23 |
| Tabla 22. Salida del procedimiento y sentencia de REG para la variable rend1718 (Rendimiento de papa t/ha) y superficie cultivada por variedad de la Tabla 21. ....                                            | 24 |



|                                                                                                                                                                                                                                                 |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabla 23. Procedimiento y sentencia de REG para datos de papa donde el rendimiento es afectado por seis variables. ....                                                                                                                         | 25 |
| Tabla 24. Salidas del procedimiento y sentencia de REG para datos de papa. ....                                                                                                                                                                 | 25 |
| Tabla 25. Procedimiento y sentencia de REG para datos de papa con selección de variables que inciden en el rendimiento. ....                                                                                                                    | 26 |
| Tabla 26. Salidas del procedimiento y sentencia de REG para datos de papa, selección de variables que inciden en el rendimiento. ....                                                                                                           | 26 |
| Tabla 27. Procedimiento y sentencia de REG para datos de presupuesto financiero en función al número de beneficiarios mujeres y condición de origen e idioma. Además incluye prueba de supuestos del análisis de regresión. ....                | 28 |
| Tabla 28. Resultados del procedimiento y sentencia de REG para datos de presupuesto financiero en función al número de beneficiarios mujeres y condición de origen e idioma. Además incluye prueba de supuestos del análisis de regresión. .... | 28 |
| Tabla 29. Datos en archivo sas de la Encuesta Datos de papa evaluado 2017-2018 en Cochabamba (OAP-MDRyT, 2018). ....                                                                                                                            | 32 |
| Tabla 30. Procedimiento y sentencia de CORRESP para la variable variedad y la percepción ¿P16. Cómo califica la cosecha del año agrícola? ambas de la campaña 2017-2018. ....                                                                   | 32 |
| Tabla 31. Salida del procedimiento y sentencia de CORRESP para la variable variedad y la percepción ¿P16. Cómo califica la cosecha del año agrícola? ambas de la campaña 2017-2018. ....                                                        | 32 |
| Tabla 32. Procedimiento y sentencia de CORRESP para la variable variedad, la percepción ¿P16. Cómo califica la cosecha del año agrícola de la campaña 2017-2018? y si aplico riego?. ....                                                       | 34 |
| Tabla 33. Salida del procedimiento y sentencia de CORRESP para variedad, la percepción ¿P16. Cómo califica la cosecha del año agrícola? y riego de la campaña 2017-2018. ....                                                                   | 34 |
| Tabla 34. Procedimiento y sentencia de DATA para el grupo de datos del gasto no alimentario dentro el hogar de varios productos (Encuesta de Hogar. INE, 2017). ....                                                                            | 37 |
| Tabla 35. Procedimiento y sentencia de CLUSTER y TREE para el gasto alimentario de varios productos dentro el hogar. ....                                                                                                                       | 37 |
| Tabla 36. Salidas del procedimiento y sentencia de CLUSTER y TREE para el gasto alimentario de varios productos dentro el hogar. ....                                                                                                           | 38 |
| Tabla 37. Procedimiento PRINCOMP y sentencia de Componentes principales para el gasto alimentario de varios productos dentro el hogar y generación de figura. ....                                                                              | 39 |
| Tabla 38. Salidas del procedimiento y sentencia de componentes principales para el gasto alimentario de varios productos dentro el hogar. ....                                                                                                  | 40 |
| Tabla 39. Procedimiento SORT y GPLOT con sentencias para graficar los clusters en función a los componentes principales 1-2 y 3-4. ....                                                                                                         | 42 |
| Tabla 40. Figuras de los clusters en función a los componentes principales 1-2 y 3-4. ....                                                                                                                                                      | 42 |
| Tabla 41. Procedimiento y sentencia de DATA para datos de producción de papa. ....                                                                                                                                                              | 44 |
| Tabla 42. Procedimiento y sentencia de FACTOR, CLUSTER, TREE, SORT, DATA y MEANS para datos de producción de papa. ....                                                                                                                         | 44 |
| Tabla 43. Salidas del procedimiento y sentencia de FACTOR, CLUSTER, TREE, SORT, DATA y MEANS, para datos de producción de papa. ....                                                                                                            | 45 |
| Tabla 44. Procedimiento LOGISTIC con sentencias para determinar el rendimiento (categorizado) como es afectada por las covariables, con y sin selección por el método Stepwise. ....                                                            | 48 |
| Tabla 45. Resultados del procedimiento LOGISTIC con sentencias para determinar el rendimiento (categorizado) como es afectada por las covariables, con y sin selección por el método Stepwise. ....                                             | 49 |



## 1. ORGANIZACIÓN DE DATOS

### 1.1. Ventanas del SAS

El programa presenta varias ventanas, ellas son **Editor**, **Log**, **Output**, **Resultados** y **Explorador**.

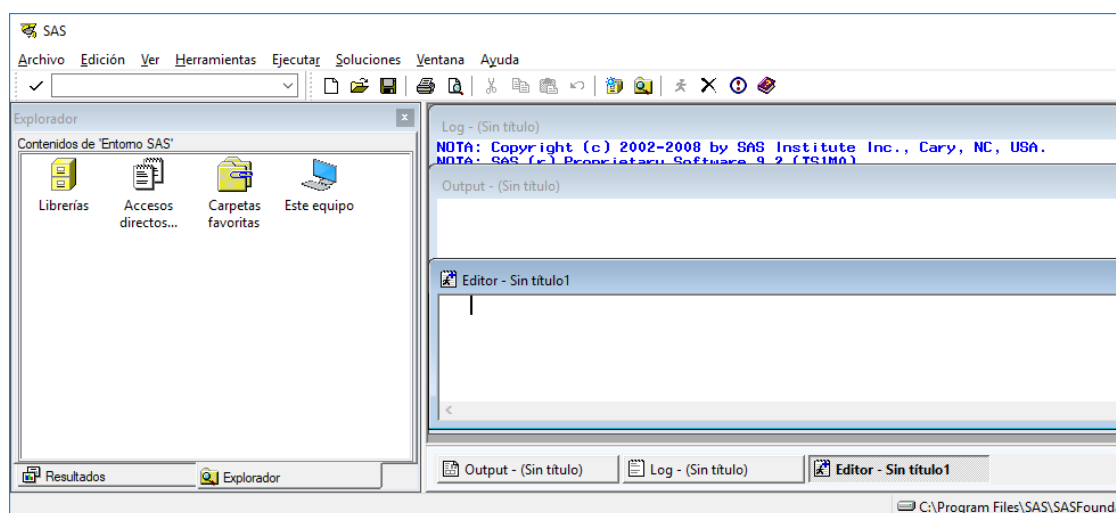


Figura 1. Vista de las ventanas de SAS

En **Editor**, se escriben los procedimientos y sentencias para de definición de la base de datos y análisis estadístico.

En **Log** visualiza el proceso de la corrida, especialmente tiempo, detalles básicos de la base de datos, los procedimientos empleados, advertencias de error y/o consideraciones.

En **Output**, expone los resultados de los procedimientos y sentencias ejecutados en Editor.

### 1.2. Introducción de datos en ventana Editor

**Utilizando archivo ascii**<sup>1</sup>. Con información parcial de precio de papa en Cochabamba en junio y julio 2018 (OAP-MDRyT, 2018 en <http://200.87.150.244:8080/birt/report/index.jsp>) (archivo: precio papa arroba Bs jun jul.xlsx), organizados en columnas las variables y en filas las observaciones, debe guardar con formato **texto delimitado por espacio** (.prn). Éste archivo recuperar a través de la ventana de Editor de SAS, escribir el procedimiento DATA y PRINT de la siguiente manera. Posteriormente guardar el archivo.

<sup>1</sup> Acrónimo ingles de American Standard Code for Information Interchange – Código Estándar Estadounidense para el Intercambio de Información)





Tabla 1. Detalles descritos sobre procedimientos y sentencias de DATA y PRINT

| Detalles                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Procedimientos y sentencias                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>data</b> Crea fichero de datos SAS</p> <p><b>input</b> Para declarar las variables del fichero (variables modalidad texto al final lleva \$)</p> <p><b>label</b> Permite nombrar una variable</p> <p><b>title1</b> Permite incluir título(s) (titlen)</p> <p><b>cards</b> Después de cards vienen los datos</p> <p>datos conformada por variables (columnas) y filas (observaciones). Un dato faltante, en su lugar un punto.</p> <p><b>;</b> finaliza una orden</p> <p><b>nodate</b> excluye la fecha del día</p> <p><b>ps</b> define el número de filas de la página</p> <p><b>proc</b> inicia un procedimiento SAS</p> <p><b>print</b> procedimiento para visualizar en la ventana Output</p> <p><b>run</b> termina la ejecución del procedimiento y/o sentencia</p> <p>Nota:<br/>           Texto azul y negrilla son procedimientos<br/>           Texto azul claro son sentencias<br/>           Texto negro son variables y nombres que uno escribe<br/>           Texto rojo expresa error</p> | <pre> data precio; input mes 1 capital \$ 7-16 bsarrb; label bsarrb=Precio mayorista de papa (Bs/arroba); title1 'Precio mayorista de papa en capitales de Bolivia, jun y jul 2018 (Bs/@)'; cards; 6 LA PAZ 25.00 6 COCHABAMBA 24.53 6 SANTA CRUZ 29.44 6 SUCRE 27.34 6 TRINIDAD 37.00 6 ORURO 27.96 6 El ALTO 29.68 6 POTOSI 29.30 7 LA PAZ 25.00 7 COCHABAMBA 26.03 7 SANTA CRUZ 28.00 7 SUCRE 26.54 7 TRINIDAD 37.12 7 ORURO 27.89 7 El ALTO 27.84 7 POTOSI 29.47 ; options nodate ps=400;  proc print data=precio; run; </pre> |



### 1.3. Ejecución de los procedimientos y sentencias

Para ejecutar el programa hacer un clic en  en el menú principal.

| Precio mayorista de papa en capitales de Bolivia, jun y jul 2018 (Bs/@) |     |            |        | 1 |
|-------------------------------------------------------------------------|-----|------------|--------|---|
| Obs                                                                     | mes | capital    | bsarrb |   |
| 1                                                                       | 6   | LA PAZ     | 25.00  |   |
| 2                                                                       | 6   | COCHABAMBA | 24.53  |   |
| 3                                                                       | 6   | SANTA CRUZ | 29.44  |   |
| 4                                                                       | 6   | SUCRE      | 27.34  |   |
| 5                                                                       | 6   | TRINIDAD   | 37.00  |   |
| 6                                                                       | 6   | ORURO      | 27.96  |   |
| 7                                                                       | 6   | EI ALTO    | 29.68  |   |
| 8                                                                       | 6   | POTOSI     | 29.30  |   |
| 9                                                                       | 7   | LA PAZ     | 25.00  |   |
| 10                                                                      | 7   | COCHABAMBA | 26.03  |   |
| 11                                                                      | 7   | SANTA CRUZ | 28.00  |   |
| 12                                                                      | 7   | SUCRE      | 26.54  |   |
| 13                                                                      | 7   | TRINIDAD   | 37.12  |   |
| 14                                                                      | 7   | ORURO      | 27.89  |   |
| 15                                                                      | 7   | EI ALTO    | 27.84  |   |
| 16                                                                      | 7   | POTOSI     | 29.47  |   |

En la ventana Log se visualiza las ordenes ejecutadas, si hubiese error esta es descrita con texto rojo.

NOTA: inicio de SAS utilizado:

tiempo real      0.88 segundos  
tiempo de cpu      0.70 segundos

```
1 data precio;
2 input mes 1 capital $ 7-16 bsarrb;
3 label bsarrb=Precio mayorista de papa (Bs/arroba);
4 title1 'Precio mayorista de papa en capitales de Bolivia, jun y jul 2018 (Bs/@) ';
5 cards;
```

NOTA: The data set WORK.PRECIO has 16 observations and 3 variables.

NOTA: Sentencia DATA ha utilizado (Tiempo de proceso total):



tiempo real 0.03 segundos  
tiempo de cpu 0.03 segundos

```
22 ;
23 options nodate ps=400;
24
25 proc print data=precio;
26 run;
```

NOTA: There were 16 observations read from the data set WORK.PRECIO.

NOTA: PROCEDIMIENTO PRINT ha utilizado (Tiempo de proceso total):

tiempo real 0.03 segundos  
tiempo de cpu 0.03 segundos

Las salidas de la ventana Output puede guardar y recuperar con Excel o Word bajo la modalidad todo tipo de archivo.

## 2. ANÁLISIS UNIVARIABLE

Los procedimientos para el análisis univariable se tiene a UNIVARIATE, MEANS para datos cuantitativos.

### 2.1. Procedimiento UNIVARIATE

Proporciona lo siguiente:

- Estadísticas descriptivas basadas en momentos (incluidas asimetría y curtosis), cuantiles o percentiles (como la mediana), tablas de frecuencia y valores extremos
- Histogramas que, opcionalmente, pueden ajustarse con curvas de densidad de probabilidad para varias distribuciones y con estimaciones de densidad de kernel
- Funciones de distribución acumulativa gráficos (gráficos cdf). Opcionalmente, estos pueden superponerse con curvas de distribución de probabilidad para varias distribuciones.
- Gráficos de cuantiles-cuantiles (gráficos Q-Q), gráficos de probabilidad y gráficos de probabilidad-probabilidad (gráficos P-P). Estas gráficas facilitan la comparación de una distribución de datos con varias distribuciones teóricas.
- Pruebas de bondad de ajuste para una variedad de distribuciones, incluida la normal.
- La posibilidad de insertar estadísticas de resumen.
- La capacidad de analizar conjuntos de datos con una variable de frecuencia.
- La capacidad de crear conjuntos de datos de salida que contienen estadísticas de resumen, intervalos de histograma y parámetros de curvas ajustadas

**Ejemplo:**



De la información del archivo **Datos de papa evaluado 2017-2018.xlsx** ('Encuesta papa Cochabamba 2017 (OAP-MDRyT, 2018)) se consideró variables y el rendimiento de papa de la campaña 2017-2018 y se expone en la Tabla 2 (archivo: **Datos de papa evaluado 2017-2018 programa.sas**).

Tabla 2. Procedimiento DATA e información de la Encuesta de papa en Cochabamba 2017

```
data papaen;
input produce 1-3 p01dep 6 p02pro 9-11 p03mun 13-17 p05prod $ 20-51 folio 58-60 varied
sup1718 rend1718 p16ccalf semt p18msem p19mcos p20fadv p20plg1 p20enf1 p23agrqm p24guan
p24barr p24zanj p24aver p24comp p24rot p24labr p24terr p24otr p25cods p26dest p27apoy
p30rieg ;
title1 'Encuesta papa Cochabamba (OAP-MDRyT, 2018)';
cards;
1 7 708 70801 ISIDRO ROJAS 169 4 12.00 4.79 2 0.38 10 2 1 98 98 1 2 2 2 1 2 2 2 2 2 2 2 1
1 2 213 21301 SILVANO MORALES ALBERTO 129 1 10.00 5.00 2 0.28 10 4 1 98 98 1 1 2 2 1 2 2 2 2 2 2 1 2 2
...
1 3 310 31001 SIXTO INTURIAS SANDOVAL 275 1 . . . . . . . . . . . . . . . . . . . . 5 2 2
1 3 303 30302 GUIDO PAUCARA 369 1 . . . . . . . . . . . . . . . . . . . . . . . . . .
1 7 709 70902 EDIBERTO ALMENDRAS 144 1 . . . . . . . . . . . . . . . . . . . . . . . . . .
;
options nodate ps=400;
```

En el procedimiento UNIVARIATE (Tabla 3), se le solicita la prueba de la distribución normal, plot de los datos PAPAEN y realizar el análisis de la variable **rend1718**. Los resultados se observan en la Tabla 4, como las medidas de posición, variación, prueba de hipótesis para media poblacional, normalidad, cuantiles, valores extremos y figuras.

Tabla 3. Procedimientos y sentencias de UNIVARIATE para la variable rend1718 (Rendimiento de papa campaña 2017-2018 (t/ha)) (OAP-MDRyT, 2018)

```
proc univariate normal plot data=papaen;
var rend1718;
run;
```

Tabla 4. Salidas del procedimiento y sentencia de Univariate de la Tabla 3.

|                                            |            |                 |            |   |
|--------------------------------------------|------------|-----------------|------------|---|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018) |            |                 |            | 1 |
| Procedimiento UNIVARIATE                   |            |                 |            |   |
| Variable: rend1718                         |            |                 |            |   |
| Momentos                                   |            |                 |            |   |
| N                                          | 574        | Sumar pesos     | 574        |   |
| Media                                      | 7.39127178 | Observ suma     | 4242.59    |   |
| Desviación std                             | 5.64211884 | Varianza        | 31.833505  |   |
| Asimetría                                  | 0.89237176 | Curtosis        | -0.0121327 |   |
| SC no corregida                            | 49598.7341 | SC corregida    | 18240.5984 |   |
| Coef. variación                            | 76.3348854 | Media error std | 0.23549749 |   |
| Medidas estadísticas básicas               |            |                 |            |   |
| Ubicación                                  |            | Variabilidad    |            |   |
| Media                                      | 7.39127    | Desviación std  | 5.64212    |   |



Fortalecimiento de la capacidad institucional en los sectores de  
desarrollo integral con coca, tráfico ilícito de drogas y  
seguridad alimentaria para una eficiente gestión del apoyo  
presupuestario sectorial – Contrato n° DCI/LA/2017/392-699



Mediana 5.63000 Varianza 31.83351  
Moda 13.80000 Rango 24.66000  
Rango intercuantil 8.26000

Tests para posición: Mu0=0

Test -Estadístico- ----P-valor-----

T de Student t 31.38578 Pr > |t| <.0001  
Signo M 287 Pr >= |M| <.0001  
Puntuación con signo S 82512.5 Pr >= |S| <.0001

Tests para normalidad

Test -Estadístico-- ----P-valor-----

Shapiro-Wilk W 0.911481 Pr < W <0.0001  
Kolmogorov-Smirnov D 0.144051 Pr > D <0.0100  
Cramer-von Mises W-Sq 2.721432 Pr > W-Sq <0.0050  
Anderson-Darling A-Sq 15.96591 Pr > A-Sq <0.0050

Cuantiles (Definición 5)

Cuantil Estimador

100% Máx 24.72  
99% 23.00  
95% 18.75  
90% 15.82  
75% Q3 11.25  
50% Mediana 5.63  
25% Q1 2.99  
10% 1.30  
5% 0.92  
1% 0.36  
0% Mín 0.06

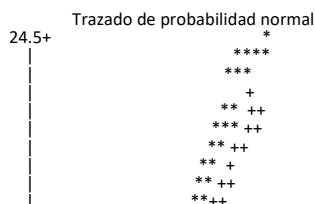
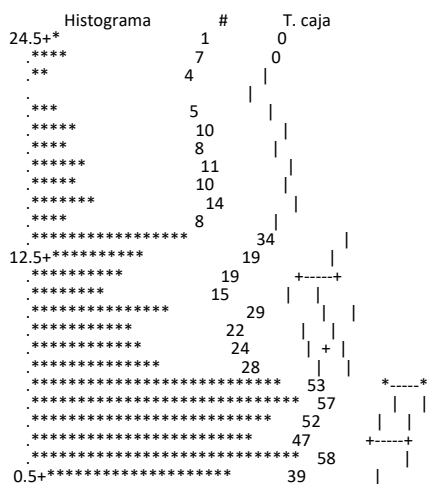
Observaciones extremas

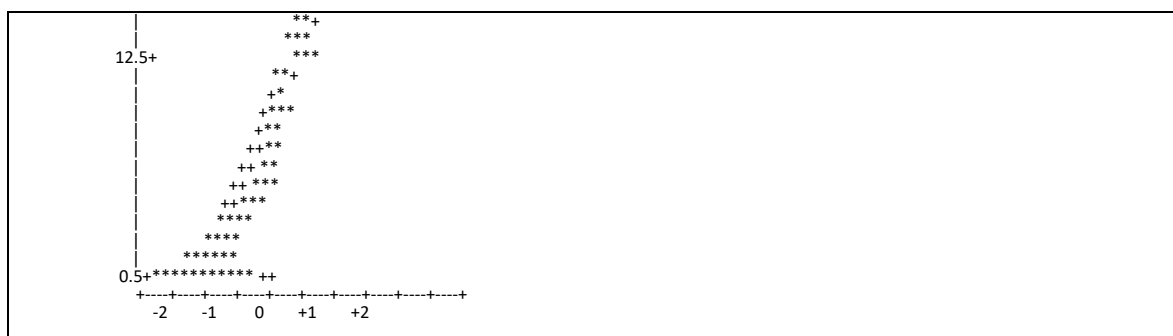
-----Inferior----- -----Superior-----

| Valor | Observación | Valor | Observación |
|-------|-------------|-------|-------------|
| 0.06  | 340         | 23.00 | 568         |
| 0.17  | 203         | 23.00 | 573         |
| 0.21  | 132         | 23.52 | 469         |
| 0.33  | 520         | 23.88 | 470         |
| 0.33  | 68          | 24.72 | 368         |

Valores ausentes

| Valor ausente | Conteo | Porcentaje de<br>Todas las<br>observaciones | Observaciones<br>ausentes |
|---------------|--------|---------------------------------------------|---------------------------|
| .             | 22     | 3.69                                        | 100.00                    |





### Comentario

El resultado anterior, la distribución normal presentó significativo a P: 0.01 según la prueba Shapiro-Wilk no tiene distribución normal. Obsérvese la media, la desviación estándar y el coeficiente de variación expresan mucha variabilidad de los datos. Las observaciones extremas también reflejan dicha variabilidad, donde el valor mínimo de 0.06 a 24.72 t/ha.

El procedimiento UNIVARIATE también realiza pruebas de distribuciones diferentes a la normal a través de la sentencia **Histogram** (Tabla 5). Los resultados se ven en la Tabla 6.

Tabla 5. Procedimientos y sentencias de UNIVARIATE para el rendimiento de papa (t/ha)

```
proc univariate normal plot data=papaen;
  histogram rend1718/ midpoints=0.00 to 26 by 2
  normal (color=red)
  lognormal (color=yellow)
  weibull (color=green)
  gamma (color=Blue) ;
  inset n mean(5.2) std='Std Dev' (5.2) skewness(5.2) / pos=ne;
  id folio;
run;
```

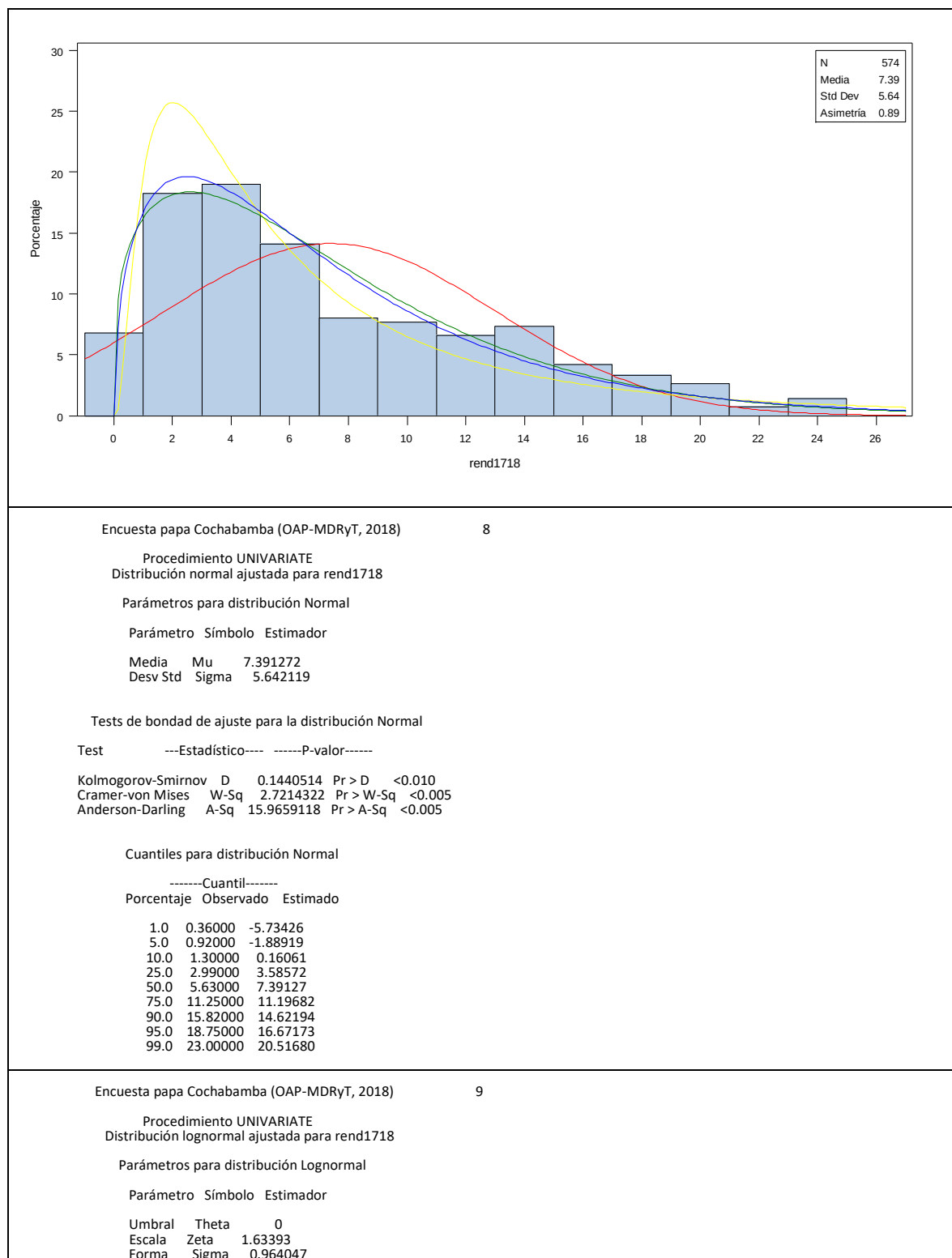
### Comentario

En la Tabla 6, la figura del histograma presenta la distribución de los datos parece aproximarse las curvas lognormal, weibull y gamma. Obsérvese las pruebas de Kolmogorov-Smirnov, Cramer-von Mises y Anderson-Darling tienen diferente nivel de rigurosidad. Así la prueba de Shapiro-Wilk presenta mayor exigencia a la Kolmogorov-Smirnov. Según estas pruebas son significativas, es decir que tienen distribuciones diferentes a la normal, lognormal, weibull y gamma a la P: 0.05.

Los rendimientos presentan variabilidad donde existe mayor concentración de datos igual y menor a 6 t/ha; menor ocurrencia de productores que presentan mayores rendimientos de papa (>16 t/ha).



Tabla 6. Salidas del procedimiento y sentencia de Univariate de la Tabla 5 (solo del rendimiento y las distribuciones).





**Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



|                                                          |                   |                   |
|----------------------------------------------------------|-------------------|-------------------|
| Media                                                    | 8.154935          |                   |
| Desv Std                                                 | 10.09683          |                   |
| Tests de bondad de ajuste para la distribución Lognormal |                   |                   |
| Test                                                     | ---Estadístico--- | -----P-valor----- |
| Kolmogorov-Smirnov                                       | D 0.07391947      | Pr > D <0.010     |
| Cramer-von Mises                                         | W-Sq 0.81509412   | Pr > W-Sq <0.005  |
| Anderson-Darling                                         | A-Sq 5.56076619   | Pr > A-Sq <0.005  |
| Cuantiles para distribución Lognormal                    |                   |                   |
|                                                          | -----Cuantil----- |                   |
| Porcentaje                                               | Observado         | Estimado          |
| 1.0                                                      | 0.36000           | 0.54402           |
| 5.0                                                      | 0.92000           | 1.04940           |
| 10.0                                                     | 1.30000           | 1.48952           |
| 25.0                                                     | 2.99000           | 2.67431           |
| 50.0                                                     | 5.63000           | 5.12397           |
| 75.0                                                     | 11.25000          | 9.81753           |
| 90.0                                                     | 15.82000          | 17.62654          |
| 95.0                                                     | 18.75000          | 25.01926          |
| 99.0                                                     | 23.00000          | 48.26170          |

|                                                        |                   |                   |
|--------------------------------------------------------|-------------------|-------------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)             |                   | 10                |
| Procedimiento UNIVARIATE                               |                   |                   |
| Distribución weibull ajustada para rend1718            |                   |                   |
| Parámetros para distribución Weibull                   |                   |                   |
| Parámetro                                              | Símbolo           | Estimador         |
| Umbral                                                 | Theta             | 0                 |
| Escala                                                 | Sigma             | 8.006482          |
| Forma                                                  | C                 | 1.300684          |
| Media                                                  |                   | 7.393817          |
| Desv Std                                               |                   | 5.732687          |
| Tests de bondad de ajuste para la distribución Weibull |                   |                   |
| Test                                                   | ---Estadístico--- | -----P-valor----- |
| Cramer-von Mises                                       | W-Sq 0.26695229   | Pr > W-Sq <0.010  |
| Anderson-Darling                                       | A-Sq 1.73039157   | Pr > A-Sq <0.010  |
| Cuantiles para distribución Weibull                    |                   |                   |
|                                                        | -----Cuantil----- |                   |
| Porcentaje                                             | Observado         | Estimado          |
| 1.0                                                    | 0.36000           | 0.23306           |
| 5.0                                                    | 0.92000           | 0.81602           |
| 10.0                                                   | 1.30000           | 1.41922           |
| 25.0                                                   | 2.99000           | 3.07213           |
| 50.0                                                   | 5.63000           | 6.04038           |
| 75.0                                                   | 11.25000          | 10.29210          |
| 90.0                                                   | 15.82000          | 15.20277          |
| 95.0                                                   | 18.75000          | 18.61186          |
| 99.0                                                   | 23.00000          | 25.90374          |

|                                                      |                   |                   |
|------------------------------------------------------|-------------------|-------------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)           |                   | 11                |
| Procedimiento UNIVARIATE                             |                   |                   |
| Distribución ajusada gamma para rend1718             |                   |                   |
| Parámetros para distribución Gamma                   |                   |                   |
| Parámetro                                            | Símbolo           | Estimador         |
| Umbral                                               | Theta             | 0                 |
| Escala                                               | Sigma             | 4.895578          |
| Forma                                                | Alfa              | 1.509785          |
| Media                                                |                   | 7.391272          |
| Desv Std                                             |                   | 6.01536           |
| Tests de bondad de ajuste para la distribución Gamma |                   |                   |
| Test                                                 | ---Estadístico--- | -----P-valor----- |
| Kolmogorov-Smirnov                                   | D 0.05160085      | Pr > D 0.002      |
| Cramer-von Mises                                     | W-Sq 0.22929019   | Pr > W-Sq 0.004   |
| Anderson-Darling                                     | A-Sq 1.66329754   | Pr > A-Sq <0.001  |
| Cuantiles para distribución Gamma                    |                   |                   |





| -----Cuantil----- |           |          |
|-------------------|-----------|----------|
| Porcentaje        | Observado | Estimado |
| 1.0               | 0.36000   | 0.28782  |
| 5.0               | 0.92000   | 0.87620  |
| 10.0              | 1.30000   | 1.45132  |
| 25.0              | 2.99000   | 3.00069  |
| 50.0              | 5.63000   | 5.83870  |
| 75.0              | 11.25000  | 10.11866 |
| 90.0              | 15.82000  | 15.37751 |
| 95.0              | 18.75000  | 19.21179 |
| 99.0              | 23.00000  | 27.86622 |

## 2.2. Procedimiento MEANS

El procedimiento MEANS proporciona herramientas de resumen de datos para calcular estadísticas descriptivas de variables de todas las observaciones y dentro de grupos de observaciones. Y proporciona lo siguiente:

- Calcula estadísticas descriptivas basadas en momentos.
- Estimaciones cuantiles, que incluye la mediana
- Calcula límites de confianza para la media.
- Identifica valores extremos
- Realiza una prueba t.

De la información de la Encuesta papa Cochabamba 2017 (OAP-MDRyT, 2018) se consideró variables y el rendimiento de papa de la campaña 2017-2018 plasmada en el archivo: **Datos de papa evaluado 2017-2018 programa.sas**, en el SAS emplear el procedimiento MEANS solicitando estadística descriptiva (Tabla 7) y los resultados se visualiza en la Tabla 8.

Tabla 7. Procedimientos y sentencias de MEANS para la variable rendimiento de papa (t/ha)

```
proc means data=papaen n mean std min max clm cv median maxdec=2;
var rend1718;
run;
```

Tabla 8. Salidas del procedimiento y sentencia de MEANS de la Tabla 7.

|                                            |       |         |              |              |               |               |
|--------------------------------------------|-------|---------|--------------|--------------|---------------|---------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018) |       |         |              |              |               | 33            |
| Procedimiento MEANS                        |       |         |              |              |               |               |
| Variable de análisis: rend1718             |       |         |              |              |               |               |
|                                            |       |         | Inferior 95% | Superior 95% |               |               |
| N                                          | Media | Dev tip | Mínimo       | Máximo       | CL para media | CL para media |
| 574                                        | 7.39  | 5.64    | 0.06         | 24.72        | 6.93          | 7.85          |
| Variable de análisis: rend1718             |       |         |              |              |               |               |
| Coeficiente                                |       |         |              |              |               |               |
| de variación                               |       |         |              |              |               |               |
| Mediana                                    |       |         |              |              |               |               |



76.33 5.63  
 ffffffffffffffffffffffff

## Comentario

También se observa variabilidad de los datos del rendimiento, desde un valor mínimo de 0.06 a 24.72 t/ha. La media poblacional del rendimiento de papa se estima valores entre los intervalos inferior y superior al 95 % de certeza y que corresponde a 6.93 a 7.85 t/ha.

## 2.3. Procedimiento FREQ

El procedimiento FREQ produce tablas de frecuencia unidireccional a n-vías y de contingencia (tabla de contingencia). Para las tablas de frecuencia unidireccional, calcula las pruebas de bondad de ajuste para proporciones iguales o proporciones nulas especificadas; también proporciona límites de confianza y pruebas para proporciones binomiales.

Para las tablas de contingencia, PROC FREQ puede calcular varias estadísticas para examinar las relaciones entre dos variables de clasificación e incluyen lo siguiente:

- Pruebas de chi-cuadrado y medidas
- Medidas de asociación
- Riesgos (proporciones binomiales) y diferencias de riesgo para tablas.
- Odds ratios y riesgos relativos para tablas.
- Pruebas de tendencia
- Pruebas y medidas de acuerdo.
- Estadísticas de Cochran-Mantel-Haenszel

## Distribución de frecuencia (binomio)

Continuando con el archivo **Datos de papa evaluado 2017-2018 programa.sas**, en el SAS emplear el procedimiento FREQ solicitando: en primera instancia la distribución de frecuencias de la Aplicación de agroquímicos (p23agrqm: 1=Si y 2=No) (Tabla 9), también se solicita la proporción binomial es intervalos de confianza; en segunda instancia también se solicita la distribución de frecuencias, la proporción binomial con la prueba de hipótesis  $p=0.75$ . Los resultados se visualizan en la Tabla 10.

Tabla 9. Procedimientos y sentencias de FREQ para la variable P23 ¿Utilizo algún agroquímico en la campaña 2016-2017?

```
proc freq data=papaen order=freq;
  tables p23agrqm / binomial(ac wilson exact) alpha=.05;
  tables p23agrqm / binomial(equiv p=.75 margin=.1) ;
run;
```



Tabla 10. Salidas del procedimiento y sentencia de FREQ de la Tabla 9.

|                                                        |            |                          |
|--------------------------------------------------------|------------|--------------------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)             |            | 36                       |
| Procedimiento FREQ                                     |            |                          |
| p23agrqm                                               | Frecuencia | Porcentaje               |
| 1                                                      | 456        | 77.82                    |
| 2                                                      | 130        | 22.18                    |
| Frecuencia de valores ausentes = 10                    |            |                          |
| Proporción binomial para p23agrqm = 1                  |            |                          |
| Proporción                                             |            | 0.7782                   |
| ASE                                                    |            | 0.0172                   |
| Tipo                                                   |            | 95% Límites de confianza |
| Wilson                                                 |            | 0.7428 0.8099            |
| Agresti-Coull                                          |            | 0.7427 0.8100            |
| Clopper-Pearson (Exacto)                               |            | 0.7423 0.8112            |
| Test de H0: Proporción = 0.5                           |            |                          |
| ASE bajo H0                                            |            | 0.0207                   |
| Z                                                      |            | 13.4669                  |
| Pr de un lado > Z                                      |            | <.0001                   |
| Pr de dos lados >  Z                                   |            | <.0001                   |
| Tamaño efectivo de la muestra = 586                    |            |                          |
| Frecuencia de valores ausentes = 10                    |            |                          |
| p23agrqm                                               | Frecuencia | Porcentaje               |
| 1                                                      | 456        | 77.82                    |
| 2                                                      | 130        | 22.18                    |
| Frecuencia de valores ausentes = 10                    |            |                          |
| Proporción binomial para p23agrqm = 1                  |            |                          |
| Proporción                                             |            | 0.7782                   |
| ASE                                                    |            | 0.0172                   |
| 95% Límite conf. inferior                              |            | 0.7445                   |
| 95% Límite conf. superior                              |            | 0.8118                   |
| Límites conf. exactos                                  |            |                          |
| 95% Límite conf. inferior                              |            | 0.7423                   |
| 95% Límite conf. superior                              |            | 0.8112                   |
| Test de H0: Proporción = 0.75                          |            |                          |
| ASE bajo H0                                            |            | 0.0179                   |
| Z                                                      |            | 1.5741                   |
| Pr de un lado > Z                                      |            | 0.0577                   |
| Pr de dos lados >  Z                                   |            | 0.1155                   |
| Análisis de equivalencia                               |            |                          |
| H0: P - p0 <= Margen inferior o >= Margen superior     |            |                          |
| Ha: Margen inferior < P - p0 < Margen superior         |            |                          |
| p0 = 0.75 Margen inferior = -0.1 Margen superior = 0.1 |            |                          |
| Proporción                                             |            | ASE (Muestra)            |
| 0.7782                                                 |            | 0.0172                   |
| Dos tests de un lado (TOST)                            |            |                          |
| Test                                                   |            | Z P-Valor                |
| Margen inferior                                        |            | 7.4668 Pr > Z <.0001     |
| Margen superior                                        |            | -4.1858 Pr < Z <.0001    |
| Total                                                  |            | <.0001                   |
| Límites de equivalencia                                |            | 90% Límites de confianza |
| 0.6500                                                 |            | 0.8500 0.7499 0.8064     |
| Tamaño efectivo de la muestra = 586                    |            |                          |
| Frecuencia de valores ausentes = 10                    |            |                          |



## Comentario

Los resultados de la Tabla 10, presenta la distribución de frecuencias, en la muestra se observa un 77.82 % de 586 encuestados. A nivel poblacional existen tres tipos de estimación de Wilson, Agresti-Coull y Clopper-Pearson; con la primera la proporción estaría de 74.28 a 80.99 %; y el test de hipótesis donde la proporción es igual a cero, la prueba es significativa, es decir que la proporción a nivel poblacional es diferente de cero. En la segunda instancia difiere la prueba de hipótesis  $p=0.75$  que fue solicitada y según la prueba (Z) es no significativa a  $P: 0.05$ , es decir la proporción a nivel muestral es similar al 75 %.

## Distribución de frecuencia (multinomial)

Con la información del archivo **Datos de papa evaluado 2017-2018 programa.sas** utilizar el proc FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? (p16ccalf: donde 1=Bueno, 2=Regular y 3=Malo) solicitando distribución de frecuencia y proporción binomial en función a la clase de mayor ocurrencia con una prueba de hipótesis de  $p=0.5$  e intervalos de confianza (Tabla 11).

Tabla 11. Procedimientos y sentencias de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018?

```
proc freq data=papaen order=freq;
  tables p16ccalf/ binomial(ac wilson exact) alpha=.05;
run;
```

Tabla 12. Salidas del procedimiento y sentencia de FREQ de la Tabla 11.

|                                            |          |            |            |                      |                      |
|--------------------------------------------|----------|------------|------------|----------------------|----------------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018) |          |            |            |                      | 37                   |
| Procedimiento FREQ                         |          |            |            |                      |                      |
|                                            | p16ccalf | Frecuencia | Porcentaje | Frecuencia acumulada | Porcentaje acumulado |
| ~~~~~                                      |          |            |            |                      |                      |
|                                            | 2        | 319        | 55.19      | 319                  | 55.19                |
|                                            | 1        | 130        | 22.49      | 449                  | 77.68                |
|                                            | 3        | 129        | 22.32      | 578                  | 100.00               |
| ~~~~~                                      |          |            |            |                      |                      |
| Frecuencia de valores ausentes = 18        |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| Proporción binomial para p16ccalf = 2      |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| Proporción 0.5519                          |          |            |            |                      |                      |
| ASE 0.0207                                 |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| Tipo 95% Límites de confianza              |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| Wilson 0.5112 0.5920                       |          |            |            |                      |                      |
| Agresti-Coull 0.5111 0.5920                |          |            |            |                      |                      |
| Clopper-Pearson (Exacto) 0.5103 0.5930     |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| Test de H0: Proporción = 0.5               |          |            |            |                      |                      |
| ~~~~~                                      |          |            |            |                      |                      |
| ASE bajo H0 0.0208                         |          |            |            |                      |                      |
| Z 2.4957                                   |          |            |            |                      |                      |
| Pr de un lado > Z 0.0063                   |          |            |            |                      |                      |
| Pr de dos lados >  Z  0.0126               |          |            |            |                      |                      |



Tamaño efectivo de la muestra = 578  
Frecuencia de valores ausentes = 18

## Comentario

Los resultados de la Tabla 12, la distribución de frecuencias con mayor ocurrencia es la clase 2 (calificación cosecha regular) con una proporción muestral de 55.19 %; los límites de confianza al 95% según Wilson sería de 51.12 a 59.20 %; y la prueba de hipótesis  $p=0.5$  es significativa a  $P: 0.05$  ( $Pr > |Z|$ ) es decir la proporción a nivel de la proporción poblacional es diferente a 50%.

## 3. ANÁLISIS BIVARIABLE

### 3.1. Entre dos variables cualitativas: tablas de contingencia y prueba de chi cuadrado.

La tabla de contingencia y la prueba de chi cuadrado tiene el objetivo de determinar si existe relación (dependencia) entre dos variables categóricas. La hipótesis nula sería que no existe relación o son independientes y la hipótesis alternativa que poseen relación o dependencia.

Se obtiene el valor de Chi cuadrado mediante:

$$\chi^2 = \sum_{i=1}^f \sum_{j=1}^c \frac{(\theta_{ij} - \xi_{ij})^2}{\xi_{ij}}$$

donde:

$\theta_{ij}$ : Frecuencia observada en la i-ésima fila, j-ésima columna

$\xi_{ij}$ : Frecuencia esperada en la i-ésima fila, j-ésima columna

f : número de filas

c : número de columnas

gl = (f - 1)(c - 1).

El programa SAS presenta la  $Pr > \chi^2$  cuadrado, con ella se puede inferir.

Empleando el archivo: **Datos de papa evaluado 2017-2018 programa.sas**, incluir el procedimiento FREQ, la sentencia tables, las variables P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? (p16ccalf) por variedad (varied) y dentro la opción adicionar la prueba de Chi cuadrado y generar en una nueva data sas las frecuencias observadas, estimadas y en porcentaje sobre el total de observaciones (Tabla 13). También se solicita la impresión de la nueva data sas.

Tabla 13. Procedimientos y sentencias de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? por variedad cultivada



```
proc freq data=papaen order=freq;
    tables p16ccalf*varied / chisq out=FreqCount outexpect sparse;
run;
proc print data=FreqCount;
run;
```

Tabla 14. Salidas del procedimiento y sentencia de FREQ para la variable P16 ¿Cómo califica la cosecha del año agrícola 2017-2018? por variedad cultivada Tabla 13.

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |  |   |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|---|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  | 1 |
| Procedimiento FREQ<br>Tabla de p16ccalf por varied                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |  |   |
| p16ccalf varied                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |  |   |
| Frecuencia,<br>Porcentaje,<br>Pct fila ,<br>Pct col , 1, 4, 6, 3, 2, 5, Total                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |  |   |
| <pre> ##### 2, 156, 83, 26, 25, 23, 6, 319 , 26.99, 14.36, 4.50, 4.33, 3.98, 1.04, 55.19 , 48.90, 26.02, 8.15, 7.84, 7.21, 1.88, , 56.12, 51.55, 54.17, 62.50, 58.97, 50.00, ##### 1, 53, 41, 13, 9, 9, 5, 130 , 9.17, 7.09, 2.25, 1.56, 1.56, 0.87, 22.49 , 40.77, 31.54, 10.00, 6.92, 6.92, 3.85, , 19.06, 25.47, 27.08, 22.50, 23.08, 41.67, ##### 3, 69, 37, 9, 6, 7, 1, 129 , 11.94, 6.40, 1.56, 1.04, 1.21, 0.17, 22.32 , 53.49, 28.68, 6.98, 4.65, 5.43, 0.78, , 24.82, 22.98, 18.75, 15.00, 17.95, 8.33, ##### Total 278 161 48 40 39 12 578 48.10 27.85 8.30 6.92 6.75 2.08 100.00 </pre> |  |   |
| Frecuencia de valores ausentes = 18                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |  |   |
| Estadísticos para la tabla de p16ccalf por varied                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |  |   |
| <pre> Estadístico DF Valor Prob ##### Chi-cuadrado 10 8.9179 0.5399 Chi-cuadrado de ratio de verosimilitud 10 9.0109 0.5311 Chi-cuadrado Mantel-Haenszel 1 3.9576 0.0467 Coeficiente Phi 0.1242 Coeficiente de contingencia 0.1233 V de Cramer 0.0878 </pre>                                                                                                                                                                                                                                                                                                                                       |  |   |
| Tamaño efectivo de la muestra = 578<br>Frecuencia de valores ausentes = 18                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |   |
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  | 2 |
| Obs p16ccalf varied COUNT EXPECTED PERCENT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |   |
| <pre> 1 . 1 6 . . 2 . 4 9 . . 3 . 6 1 . . 4 . 3 1 . . 5 . 2 0 . . 6 . 5 1 . . 7 2 1 156 153.429 26.9896 8 2 4 83 88.856 14.3599 9 2 6 26 26.491 4.4983 10 2 3 25 22.076 4.3253 11 2 2 23 21.524 3.9792 12 2 5 6 6.623 1.0381 13 1 1 53 62.526 9.1696 14 1 4 41 36.211 7.0934 15 1 6 13 10.796 2.2491 16 1 3 9 8.997 1.5571 17 1 2 9 8.772 1.5571 18 1 5 5 2.699 0.8651 19 3 1 69 62.045 11.9377 20 3 4 37 35.933 6.4014 21 3 6 9 10.713 1.5571 </pre>                                                                                                                                              |  |   |



|    |   |   |   |       |        |
|----|---|---|---|-------|--------|
| 22 | 3 | 3 | 6 | 8.927 | 1.0381 |
| 23 | 3 | 2 | 7 | 8.704 | 1.2111 |
| 24 | 3 | 5 | 1 | 2.678 | 0.1730 |

En la Tabla 14, la prueba de Chi cuadrado la probabilidad es mayor a 0.05 por lo cual se asume la declaración sobre la cosecha 2017-2018 es independiente a las variedades, es decir no existe relación alguna.

### 3.2. Para una variable cuantitativa y otra cualitativa: análisis de varianza.

#### *Análisis de varianza.*

La técnica del análisis de varianza se emplea para realizar comparaciones a más de dos muestras. Entre las aplicaciones están Diseños Experimentales y Muestrales, algunas diferencias se exponen en la Tabla 15.

Tabla 15. Al mismo tiempo se debe diferenciar las características del análisis de varianza en el muestreo y diseño experimental (Leon y Quiroz, 1994):

| Característica                  | Muestreo                                | Diseño Experimental                                   |
|---------------------------------|-----------------------------------------|-------------------------------------------------------|
| Unidad básica                   | Unidad de muestreo                      | Unidad Experimental                                   |
| Tratamientos                    | No se aplica a las unidades de muestreo | Se aplica a las unidades experimentales               |
| Control de fuentes de variación | Estratificación                         | Bloques                                               |
| Variabilidad                    | Se mide la variación natural            | Se mide la variabilidad inducida por los tratamientos |

#### *Procedimientos para el análisis de varianza en SAS*

ANOVA: Análisis de varianza (datos balanceados y factores fijos)

GLM: Modelo lineal general (factores fijos y aleatorios)

MIXED: Modelo lineal mixto (factores fijos y mixtos)

GENMOD: Modelo lineal generalizado (factores fijos y otras distribuciones)

GLIMMIX: Modelo lineal generalizado mixto (factores fijos, aleatorios y otras distribuciones)

GLMSELECT: Modelo lineal general selección (selección de variables similar)

El análisis de varianza de clasificación de una vía tiene el objetivo en función a una variable dependiente determinar diferencias debido a clases de la variable categórica (independiente).

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

En el SAS el modelo sería:



variable dependiente = variable independiente

el residuo tiene el supuesto de  $\varepsilon_{ij} \sim DNII(0, \sigma_e^2)$

Con los datos del archivo **Datos de papa evaluado 2017-2018 programa.sas** y el procedimiento proc GLM generar el análisis de varianza de clasificación de una vía (varied), este efecto será sobre el rendimiento de papa (rend1718) (Tabla 16). Así mismo se solicita generar la prueba de homogeneidad de varianzas, Duncan, diferencias de promedios por mínimos cuadrados con la distribución de t de Student y generar un nuevo data SAS con los residuos. También con el procedimiento UNIVARIATE se pide verificar si los residuos poseen distribución normal.

Tabla 16. Procedimientos y sentencias de GLM para la variable rend1718 (Rendimiento de papa t/ha) y variedad cultivada

```
proc glm data=papaen ;
class varied;
model rend1718= varied/ss3;
means varied/hovtest=obrien duncan;
lsmeans varied / pdiff;
output out=pa01 r=rrend1718;
run;

proc univariate data=pa01 normal plot;
var rrend1718;
run;
```

### Comentario

El análisis de varianza (Tabla 17) presenta diferencias significativas a P: 0.05 debido a las variedades, es decir al menos dos de ellas difieren. Para determinar dichas diferencias se recurrió a la Prueba Duncan a P: 0.05 donde se registró que todas las variedades son similares en cuanto al rendimiento. Por el contrario las diferencias de los promedios por mínimos cuadrados y la prueba de t de Student es mayor el rendimiento de papa y difieren la variedad 4 (desire) a la variedad 1 (Waycha) y 3 (Sani imilla) a la P: 0.05, las otras son intermedias.

Nota: La prueba de homogeneidad de varianzas es no significativa a P: 0.05 es decir las varianzas son homogéneas. Y la prueba de normalidad Shapiro-Wilk es significativa a P: 0.01 traduciéndose que los residuos tiene distribución diferente a la normal.

Tabla 17. Salidas del procedimiento y sentencia de GLM para la variable rend1718 (Rendimiento de papa t/ha) y variedad cultivada.





|                                                                                                   |         |                         |                      |         |        |          |
|---------------------------------------------------------------------------------------------------|---------|-------------------------|----------------------|---------|--------|----------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                        |         |                         |                      |         |        | 77       |
| Procedimiento GLM                                                                                 |         |                         |                      |         |        |          |
| Información de nivel de clase                                                                     |         |                         |                      |         |        |          |
| Clase                                                                                             | Niveles | Valores                 |                      |         |        |          |
| varied                                                                                            | 6       | 1                       | 2                    | 3       | 4      | 5        |
| Número de observaciones leídas                                                                    |         |                         |                      |         |        | 596      |
| Número de observaciones usadas                                                                    |         |                         |                      |         |        | 574      |
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                        |         |                         |                      |         |        | 78       |
| Procedimiento GLM                                                                                 |         |                         |                      |         |        |          |
| Variable dependiente: rend1718                                                                    |         |                         |                      |         |        |          |
| Fuente                                                                                            | DF      | Suma de cuadrados       | Cuadrado de la media | F-Valor | Pr > F |          |
| Modelo                                                                                            | 5       | 420.85318               | 84.17064             | 2.68    | 0.0208 |          |
| Error                                                                                             | 568     | 17819.74519             | 31.37279             |         |        |          |
| Total corregido                                                                                   | 573     | 18240.59837             |                      |         |        |          |
| R-cuadrado                                                                                        |         |                         |                      |         |        |          |
| Coef Var                                                                                          |         |                         |                      |         |        |          |
| Raíz MSE                                                                                          |         |                         |                      |         |        |          |
| rend1718 Media                                                                                    |         |                         |                      |         |        |          |
| 0.023072                                                                                          |         |                         |                      |         |        |          |
| 75.78049                                                                                          |         |                         |                      |         |        |          |
| 5.601142                                                                                          |         |                         |                      |         |        |          |
| 7.391272                                                                                          |         |                         |                      |         |        |          |
| Fuente                                                                                            | DF      | Cuadrado de Tipo III SS | Cuadrado de la media | F-Valor | Pr > F |          |
| varied                                                                                            | 5       | 420.8531776             | 84.1706355           | 2.68    | 0.0208 |          |
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                        |         |                         |                      |         |        | 79       |
| Procedimiento GLM                                                                                 |         |                         |                      |         |        |          |
| Test de O'Brien para homogeneidad de la varianza rend1718                                         |         |                         |                      |         |        |          |
| ANOVA de variable de extensión de O'Brien, W = 0.5                                                |         |                         |                      |         |        |          |
| Fuente                                                                                            | DF      | Suma de cuadrados       | Cuadrado de la media | F-Valor | Pr > F |          |
| varied                                                                                            | 5       | 9954.5                  | 1990.9               | 0.93    | 0.4595 |          |
| Error                                                                                             | 568     | 1212964                 | 2135.5               |         |        |          |
| Prueba del rango múltiple de Duncan para rend1718                                                 |         |                         |                      |         |        |          |
| NOTA: This test controls the Type I comparisonwise error rate, not the experimentwise error rate. |         |                         |                      |         |        |          |
| Alpha                                                                                             |         |                         |                      |         |        | 0.05     |
| Error Degrees of Freedom                                                                          |         |                         |                      |         |        | 568      |
| Error de cuadrado medio                                                                           |         |                         |                      |         |        | 31.37279 |
| Media armónica de tamaño de celdas                                                                |         |                         |                      |         |        | 36.43119 |
| NOTA: Los tamaños de las celdas no son iguales.                                                   |         |                         |                      |         |        |          |
| Número de medias                                                                                  | 2       | 3                       | 4                    | 5       | 6      |          |
| Rango crítico                                                                                     | 2.578   | 2.714                   | 2.805                | 2.872   | 2.925  |          |
| Medias con la misma letra no son significativamente diferentes.                                   |         |                         |                      |         |        |          |
| Duncan Agrupamiento                                                                               | Media   | N                       | varied               |         |        |          |
| A                                                                                                 | 8.711   | 160                     | 4                    |         |        |          |
| A                                                                                                 | 7.829   | 12                      | 5                    |         |        |          |
| A                                                                                                 | 7.371   | 39                      | 2                    |         |        |          |
| A                                                                                                 | 6.998   | 48                      | 6                    |         |        |          |
| A                                                                                                 | 6.834   | 275                     | 1                    |         |        |          |
| A                                                                                                 | 6.303   | 40                      | 3                    |         |        |          |
| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                                                        |         |                         |                      |         |        | 81       |
| Procedimiento GLM                                                                                 |         |                         |                      |         |        |          |
| Medias de cuadrados mínimos                                                                       |         |                         |                      |         |        |          |
| rend1718                                                                                          |         |                         |                      |         |        | Número   |



**Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



| varied | LSMEAN     | LSMEAN |
|--------|------------|--------|
| 1      | 6.83447273 | 1      |
| 2      | 7.37076923 | 2      |
| 3      | 6.30325000 | 3      |
| 4      | 8.71056250 | 4      |
| 5      | 7.82916667 | 5      |
| 6      | 6.99750000 | 6      |

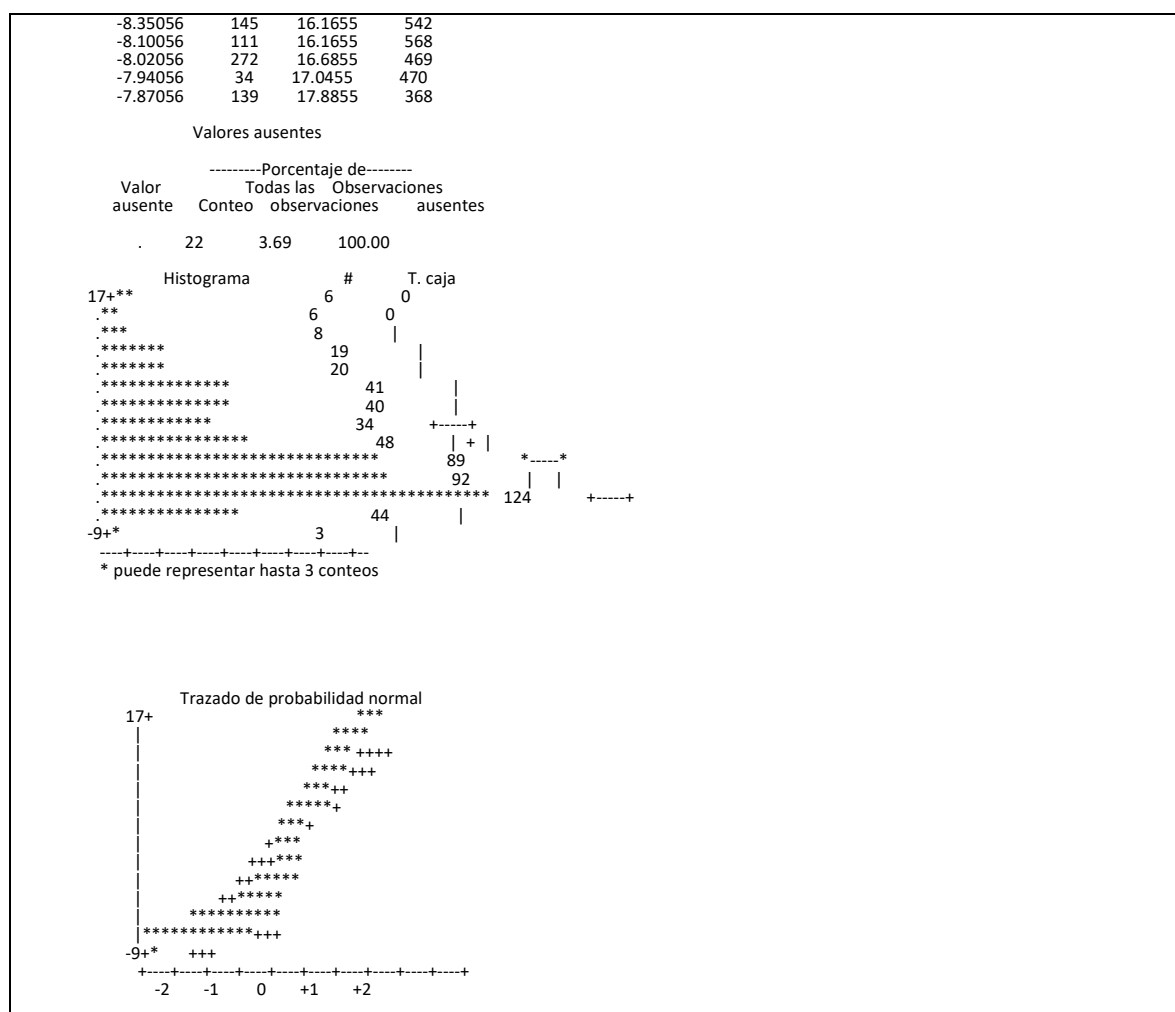
Medias de cuadrados mínimos para el efecto varied  
Pr > |t| para H0: MediaLS(i)=MediaLSn(j)

Variable dependiente: rend1718

| i/j | 1      | 2      | 3      | 4      | 5      | 6      |
|-----|--------|--------|--------|--------|--------|--------|
| 1   |        | 0.5760 | 0.5754 | 0.0008 | 0.5473 | 0.8525 |
| 2   | 0.5760 |        | 0.3974 | 0.1810 | 0.8043 | 0.7573 |
| 3   | 0.5754 | 0.3974 |        | 0.0154 | 0.4082 | 0.5628 |
| 4   | 0.0008 | 0.1810 | 0.0154 |        | 0.5993 | 0.0636 |
| 5   | 0.5473 | 0.8043 | 0.4082 | 0.5993 |        | 0.6457 |
| 6   | 0.8525 | 0.7573 | 0.5628 | 0.0636 | 0.6457 |        |

NOTA: Para asegurar un nivel de protección completo, sólo se deben usar probabilidades asociadas con comparaciones preplanificadas.

|                                            |                    |                            |
|--------------------------------------------|--------------------|----------------------------|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018) |                    | 87                         |
| Procedimiento UNIVARIATE                   |                    |                            |
| Variable: rend1718                         |                    |                            |
| Momentos                                   |                    |                            |
| N                                          | 574                | Sumar pesos 574            |
| Media                                      | 0                  | Observ suma 0              |
| Desviación std                             | 5.57665059         | Varianza 31.0990318        |
| Asimetría                                  | 0.92985996         | Curtosis 0.14549754        |
| SC no corregida                            | 17819.7452         | SC corregida 17819.7452    |
| Coef. variación                            | .                  | Media error std 0.2327649  |
| Medidas estadísticas básicas               |                    |                            |
| Ubicación                                  |                    | Variabilidad               |
| Media                                      | 0.00000            | Desviación std 5.57665     |
| Mediana                                    | -1.67886           | Varianza 31.09903          |
| Moda                                       | -1.83447           | Rango 26.23609             |
|                                            |                    | Rango intercuantil 8.09630 |
| Tests para posición: Mu0=0                 |                    |                            |
| Test                                       | -Estadístico-      | -----P-valor-----          |
| T de Student                               | t 0                | Pr >  t  1.0000            |
| Signo                                      | M -65              | Pr >=  M  <.0001           |
| Puntuación con signo                       | S -7114.5          | Pr >=  S  0.0735           |
| Tests para normalidad                      |                    |                            |
| Test                                       | -Estadístico--     | -----P-valor-----          |
| Shapiro-Wilk                               | W 0.914792         | Pr < W <0.0001             |
| Kolmogorov-Smirnov                         | D 0.12564          | Pr > D <0.0100             |
| Cramer-von Mises                           | W-Sq 2.793017      | Pr > W-Sq <0.0050          |
| Anderson-Darling                           | A-Sq 16.01219      | Pr > A-Sq <0.0050          |
| Cuantiles (Definición 5)                   |                    |                            |
| Cuantil                                    | Estimador          |                            |
| 100% Máx                                   | 17.88553           |                            |
| 99%                                        | 16.00250           |                            |
| 95%                                        | 10.83944           |                            |
| 90%                                        | 8.16553            |                            |
| 75% Q3                                     | 3.72553            |                            |
| 50% Mediana                                | -1.67886           |                            |
| 25% Q1                                     | -4.37077           |                            |
| 10%                                        | -5.84447           |                            |
| 5%                                         | -6.37447           |                            |
| 1%                                         | -7.69056           |                            |
| 0% Mín                                     | -8.35056           |                            |
| Observaciones extremas                     |                    |                            |
| -----Inferior-----                         | -----Superior----- |                            |
| Valor Observación                          | Valor Observación  |                            |



### 3.3. Entre dos variables cuantitativas: correlación y regresión lineal simple.

#### *Análisis de correlación*

El procedimiento CORR calcula los coeficientes de correlación de Pearson, tres medidas no paramétricas de asociación y las probabilidades asociadas con estas estadísticas, estas son:

- Correlación producto-momento de Pearson.
- Correlación de orden de rango de Spearman
- El coeficiente tau-b de Kendall
- La medida de dependencia de Hoeffding,
- Correlación parcial de Pearson, Spearman y Kendall

El coeficiente de correlación simple (Pearson) está dada por:



$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})/(n - 1)}{\sqrt{\sum(X - \bar{X})^2/(n - 1)} \sqrt{\sum(Y - \bar{Y})^2/(n - 1)}}$$

donde  $r$  varía entre  $-1$  a  $+1$

### Ejemplo

Empleando los datos del archivo **Datos de papa evaluado 2017-2018 programa.sas** el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1 se procesara con el proc CORR (Tabla 18).

Tabla 18. Procedimiento y sentencia de CORR (análisis de correlación) para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1.

```
proc corr data=papaen;
var rend1718 sup1718 semt p18msem p19mcos p20plg1 p20enf1 ;
run;
```

Tabla 19. Salida del procedimiento y sentencia de CORR para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1.

| Encuesta papa Cochabamba (OAP-MDRyT, 2018)                         |                  |          |          |           |          |          | 1        |
|--------------------------------------------------------------------|------------------|----------|----------|-----------|----------|----------|----------|
| Procedimiento CORR                                                 |                  |          |          |           |          |          |          |
| 7 Variables: rend1718 sup1718 semt p18msem p19mcos p20plg1 p20enf1 |                  |          |          |           |          |          |          |
| Estadísticos simples                                               |                  |          |          |           |          |          |          |
| Variable                                                           | N                | Media    | Dev tip  | Suma      | Mínimo   | Máximo   |          |
| rend1718                                                           | 574              | 7.39127  | 5.64212  | 4243      | 0.06000  | 24.72000 |          |
| sup1718                                                            | 586              | 0.94867  | 1.31881  | 555.92000 | 0.01000  | 12.00000 |          |
| semt                                                               | 584              | 1.07642  | 0.71353  | 628.63000 | 0        | 3.45000  |          |
| p18msem                                                            | 586              | 9.70819  | 4.23722  | 5689      | 1.00000  | 99.00000 |          |
| p19mcos                                                            | 586              | 5.01365  | 8.05769  | 2938      | 1.00000  | 99.00000 |          |
| p20plg1                                                            | 525              | 56.84762 | 40.24002 | 29845     | 2.00000  | 99.00000 |          |
| p20enf1                                                            | 525              | 70.49524 | 34.37299 | 37010     | 2.00000  | 99.00000 |          |
| Coeficientes de correlación Pearson                                |                  |          |          |           |          |          |          |
| Prob >  r  suponiendo H0: Rho=0                                    |                  |          |          |           |          |          |          |
| Número de observaciones                                            |                  |          |          |           |          |          |          |
|                                                                    | rend1718         | sup1718  | semt     | p18msem   | p19mcos  | p20plg1  | p20enf1  |
| rend1718                                                           | 1.00000          | -0.12200 | 0.39953  | -0.04710  | 0.09982  | -0.01466 | -0.08991 |
|                                                                    | 0.0034           | <.0001   | 0.2599   | 0.0167    | 0.7402   | 0.0416   |          |
|                                                                    | 574              | 574      | 572      | 574       | 574      | 514      | 514      |
| sup1718                                                            | -0.12200         | 1.00000  | -0.04196 | 0.00087   | -0.01456 | 0.08080  | -0.02382 |
|                                                                    | <b>0.0034</b>    | 0.3114   | 0.9833   | 0.7250    | 0.0643   | 0.5861   |          |
|                                                                    | 574              | 586      | 584      | 586       | 525      | 525      |          |
| semt                                                               | 0.39953          | -0.04196 | 1.00000  | -0.01187  | 0.06567  | 0.02629  | -0.11381 |
|                                                                    | <b>&lt;.0001</b> | 0.3114   | 0.7748   | 0.1129    | 0.5486   | 0.0092   |          |
|                                                                    | 572              | 584      | 584      | 584       | 523      | 523      |          |
| p18msem                                                            | -0.04710         | 0.00087  | -0.01187 | 1.00000   | 0.32235  | -0.03549 | 0.03502  |
|                                                                    | 0.2599           | 0.9833   | 0.7748   | <.0001    | 0.4170   | 0.4233   |          |
|                                                                    | 574              | 586      | 584      | 586       | 525      | 525      |          |



|         |               |          |               |                  |               |                  |         |
|---------|---------------|----------|---------------|------------------|---------------|------------------|---------|
| p19mcos | 0.09982       | -0.01456 | 0.06567       | 0.32235          | 1.00000       | 0.12979          | 0.04889 |
|         | <b>0.0167</b> | 0.7250   | 0.1129        | <b>&lt;.0001</b> | 0.0029        | 0.2635           |         |
|         | 574           | 586      | 584           | 586              | 525           | 525              |         |
| p20plg1 | -0.01466      | 0.08080  | 0.02629       | -0.03549         | 0.12979       | 1.00000          | 0.33369 |
|         | 0.7402        | 0.0643   | 0.5486        | 0.4170           | <b>0.0029</b> | <b>&lt;.0001</b> |         |
|         | 514           | 525      | 523           | 525              | 525           | 525              |         |
| p20enf1 | -0.08991      | -0.02382 | -0.11381      | 0.03502          | 0.04889       | 0.33369          | 1.00000 |
|         | <b>0.0416</b> | 0.5861   | <b>0.0092</b> | 0.4233           | 0.2635        | <b>&lt;.0001</b> |         |
|         | 514           | 525      | 523           | 525              | 525           | 525              |         |

## Comentario

Los resultados (Tabla 19) presenta dos partes, la primera una estadística descriptiva de todas las variables y la segunda la matriz con los coeficientes de correlación de Pearson (lineal) con su probabilidad y número de observaciones.

Referente a la correlación del rendimiento con la superficie sembrada es significativas a P: 0.01 deduciéndose que a un incremento de la superficie el rendimiento disminuye, la correlación del rendimiento y semilla empleada también es significativa a P: 0.01 y en forma positiva; este comportamiento también se observa la correlación con el mes de cosecha a P: 0.05 es decir a medida que se retrasa la cosecha aumentaría el rendimiento; ocurre una disminución del rendimiento cuando incrementa la ocurrencia de enfermedad a P: 0.05. Otras correlaciones que tienen sentido lógica donde a cosecha tardía la ocurrencia de plagas aumenta y al incremento de plagas aumentaría la enfermedad a P: 0.01.

## Ejercicio

El análisis de correlación por variedad se la obtiene incluyendo la sentencia BY. Previamente debe ordenarse los datos por variedad con el proc SORT (aplicar en sas los procedimientos y sentencias de la Tabla 20).

Tabla 20. Procedimiento y sentencia de SORT y CORR para el rendimiento de papa, superficie, semilla empleada, mes de siembra, mes de cosecha, efecto de plaga1 y efecto de enfermedad1 por variedad.

```
proc sort data=papaen; by varied; run;

proc corr data=papaen;
var rend1718 sup1718 semt p18msem p19mcos p20plg1 p20enf1 ;
by varied;
run;
```



### Regresión lineal

El procedimiento REG es uno de los muchos procedimientos de regresión en el sistema SAS. Es un procedimiento de propósito general para la regresión.

Otros procedimientos SAS/STAT que realizan al menos un tipo de análisis de regresión son los procedimientos CATMOD, GENMOD, GLM, LOGÍSTIC, MIXED, NLIN, ORTHOREG, PROBIT, RSREG y TRANSREG. Los procedimientos SAS/ETS están especializados para aplicaciones en series de tiempo o sistemas simultáneos.

El procedimiento REG proporciona las siguientes capacidades:

- Múltiples declaraciones de MODELO
- Nueve métodos de selección de modelo
- Cambios interactivos tanto en el modelo como en los datos utilizados para ajustar el modelo.
- Restricciones de igualdad lineal en los parámetros.
- Ensayos de hipótesis lineales e hipótesis multivariadas.
- Diagnóstico de colinealidad
- Valores predichos, residuales, residuales estudiados, límites de confianza y estadísticas de influencia
- Correlación o entrada de producto cruzado
- Estadísticas solicitadas disponibles para salida a través de conjuntos de datos de salida

### Ejemplo

Con la información del archivo **Datos de papa evaluado 2017-2018 programa.sas** y en la Tabla 21, el procedimiento REG, se plantea la regresión lineal del rendimiento en función a la superficie, se incluyó en opción límites de confianza de los coeficientes de regresión (clb) y la prueba de D de Durbin-Watson (dwprob). Además se pide datos de los residuos (r) y estimados (p).

Con las sentencias de grafico (goptions) y el proc GPLOT se le solicita graficar los valores estimados y observados del rendimiento en el eje de las ordenadas y en las abscisas la superficie cultivada.

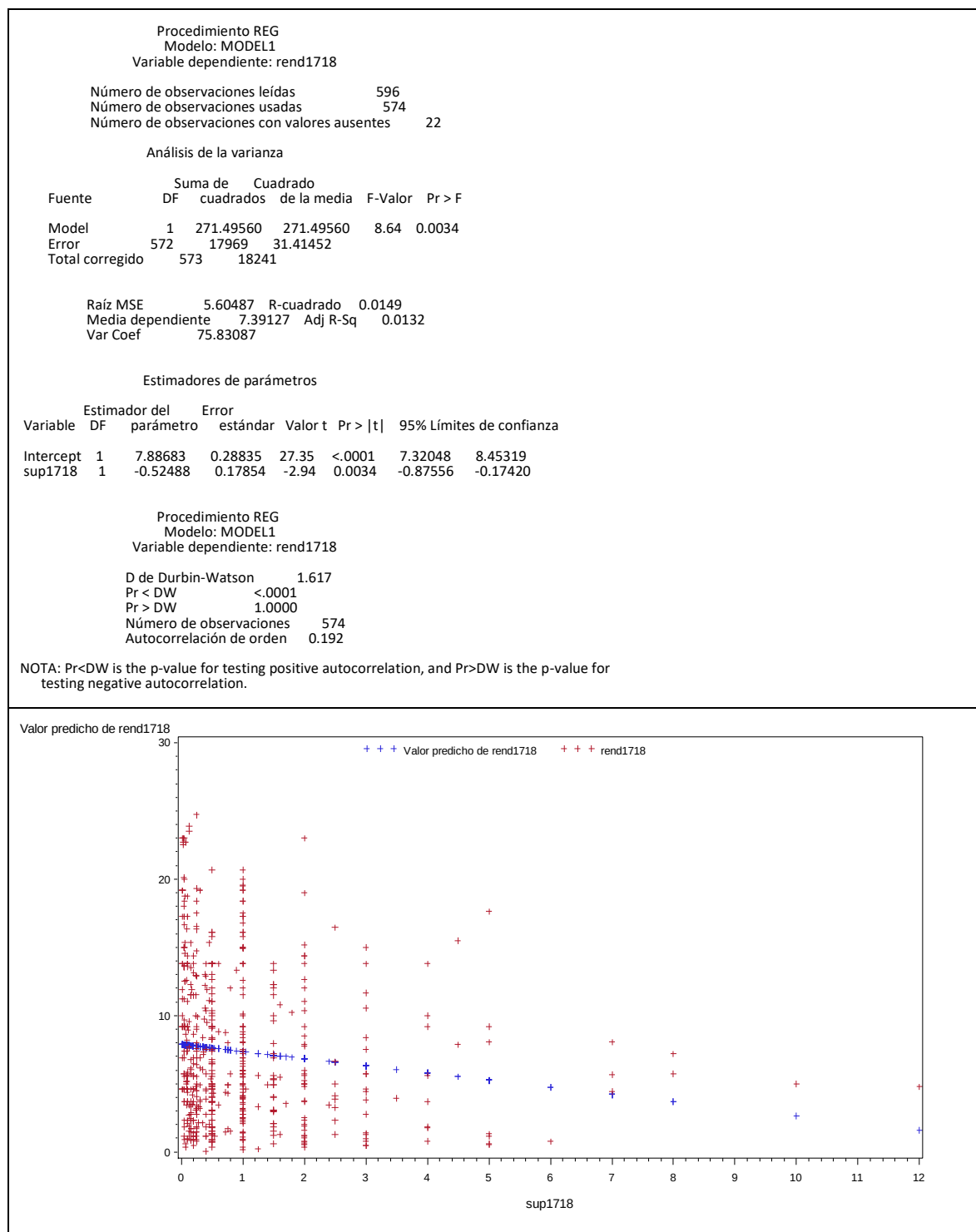
Tabla 21. Procedimiento y sentencia de REG para la variable rend1718 (Rendimiento de papa t/ha) y superficie cultivada. Y gráfica de dispersión de valores observados y estimados sobre la superficie.

```
proc reg data=papaen ;  
model rend1718 = sup1718 / clb dwprob ;  
    output out=out r=r p=p;  
run; quit;  
  
goptions reset=all border;  
legend1 label=none  
    position=(top center inside)  
    mode=share;
```



```
proc gplot data=out;
plot p*sup1718 rend1718*sup1718 / overlay legend=legend1;
run; quit;
```

Tabla 22. Salida del procedimiento y sentencia de REG para la variable rend1718 (Rendimiento de papa t/ha) y superficie cultivada por variedad de la Tabla 21.





## Comentario

Los resultados (Tabla 22) el número de observaciones utilizadas son 574; el análisis de varianza el modelo debido a la regresión es significativa a  $P: 0.01$ ; referente al coeficiente de determinación  $R^2$  expresado en porcentaje fue 1.49% es decir la variación explicada del rendimiento es causada por la superficie; así mismo, los coeficientes intercepto y regresión de la superficie sembrada según la prueba en base a la distribución de  $t$  de Student son significativas, deduciéndose a nivel poblacional son diferente de cero a  $P: 0.01$  y los límites de confianza permiten visualizar por ejemplo el coeficiente de regresión por unidad de superficie el rendimiento disminuye en  $-0.87556$  a  $-0.1742$  t/ha a la  $P: 0.05$ . Nótese el coeficiente  $D$  de Durbin-Watson según la  $Pr < DW$  es significativa a  $P: 0.01$  deduciéndose la autocorrelación positiva.

## 4. ANÁLISIS MULTIVARIABLE

### 4.1. Análisis de regresión múltiple

Con la información del archivo **Datos de papa evaluado 2017-2018 programa.sas** y en la Tabla 23, se expone el procedimiento REG considerando el análisis de regresión lineal múltiple del rendimiento en función a la superficie, semilla empleada, mes de siembra, mes de cosecha incidencia de plaga y enfermedad; así mismo se incluyó en opción límites de confianza de los coeficientes de regresión (clb) y la prueba de  $D$  de Durbin-Watson (dwprob).

Tabla 23. Procedimiento y sentencia de REG para datos de papa donde el rendimiento es afectado por seis variables.

```
options ls=110;
proc reg data=papaen;
model rend1718 = sup1718 semt p18mse p19mcos p20plg1 p20enf1 /clb dwprob ;
run;
```

Tabla 24. Salidas del procedimiento y sentencia de REG para datos de papa.

|                                              |               |                       |             |         |        |
|----------------------------------------------|---------------|-----------------------|-------------|---------|--------|
| Procedimiento REG                            |               |                       |             |         |        |
| Modelo: MODEL1                               |               |                       |             |         |        |
| Variable dependiente: rend1718               |               |                       |             |         |        |
| Número de observaciones leídas               |               | 596                   |             |         |        |
| Número de observaciones usadas               |               | 512                   |             |         |        |
| Número de observaciones con valores ausentes |               | 84                    |             |         |        |
| Análisis de la varianza                      |               |                       |             |         |        |
| Fuente                                       | Suma de<br>DF | Cuadrado<br>cuadrados | de la media | F-Valor | Pr > F |
| Model                                        | 6             | 3471.95732            | 578.65955   | 23.96   | <.0001 |
| Error                                        | 505           | 12195                 | 24.14948    |         |        |
| Total corregido                              | 511           | 15667                 |             |         |        |





|                                                                                                      |                  |           |                |         |         |                          |          |
|------------------------------------------------------------------------------------------------------|------------------|-----------|----------------|---------|---------|--------------------------|----------|
| Raíz MSE 4.91421 R-cuadrado 0.2216<br>Media dependiente 7.08111 Adj R-Sq 0.2124<br>Var Coef 69.39887 |                  |           |                |         |         |                          |          |
| Estimadores de parámetros                                                                            |                  |           |                |         |         |                          |          |
| Variable                                                                                             | Estimador del DF | parámetro | Error estándar | Valor t | Pr >  t | 95% Límites de confianza |          |
| Intercept                                                                                            | 1                | 5.56824   | 0.82434        | 6.75    | <.0001  | 3.94868                  | 7.18780  |
| sup1718                                                                                              | 1                | -0.25797  | 0.16789        | -1.54   | 0.1250  | -0.58781                 | 0.07187  |
| semt                                                                                                 | 1                | 3.18453   | 0.30438        | 10.46   | <.0001  | 2.58653                  | 3.78254  |
| p18msem                                                                                              | 1                | -0.17117  | 0.05639        | -3.04   | 0.0025  | -0.28196                 | -0.06039 |
| p19mcos                                                                                              | 1                | 0.16887   | 0.04059        | 4.16    | <.0001  | 0.08911                  | 0.24862  |
| p20plg1                                                                                              | 1                | -0.00461  | 0.00586        | -0.79   | 0.4322  | -0.01612                 | 0.00691  |
| p20enf1                                                                                              | 1                | -0.00793  | 0.00678        | -1.17   | 0.2432  | -0.02125                 | 0.00540  |

|                                |        |
|--------------------------------|--------|
| Procedimiento REG              |        |
| Modelo: MODEL1                 |        |
| Variable dependiente: rend1718 |        |
| D de Durbin-Watson             | 1.771  |
| Pr < DW                        | 0.0036 |
| Pr > DW                        | 0.9964 |
| Número de observaciones        | 512    |
| Autocorrelación de orden       | 0.112  |

NOTA: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

## Comentario

Los resultados de la Tabla 24, presenta 512 observaciones empleadas; en el análisis de varianza la fuente de variación debido al modelo es significativa a P: 0.01 y el coeficiente de determinación registro un 22.16 % deduciéndose esta variación explicada del rendimiento es debida a las seis variables. Los coeficientes de regresión significativas son semilla empleada (semt), mes de siembra (p18msem) y mes cosecha (p19mcos) a P: 0.01, es decir que poseen coeficientes de regresión a nivel poblacional diferente a 0 y los límites de confianza también lo respaldan.

El proc REG, permite realizar selección de variables que poseen incidencia según el nivel de probabilidad con el método Stepwise, en ese sentido en la Tabla 25 se expone las sentencias y opciones; en la Tabla 26 se expone los resultados.

Tabla 25. Procedimiento y sentencia de REG para datos de papa con selección de variables que inciden en el rendimiento.

```
proc reg data=papaen ;
model rend1718 = sup1718 semt p18msem p19mcos p20plg1 p20enf1 /selection=
stepwise sle=0.10;
run;
quit;
```

Tabla 26. Salidas del procedimiento y sentencia de REG para datos de papa, selección de variables que inciden en el rendimiento.

Procedimiento REG  
Modelo: MODEL1



Variable dependiente: rend1718

Número de observaciones leídas 596  
Número de observaciones usadas 512  
Número de observaciones con valores ausentes 84

Selección escalonada: Paso 1

Variable semt introducida: R-cuadrado = 0.1866 y C(p) = 19.6932

Análisis de la varianza

| Fuente          | Suma de<br>DF | Cuadrado<br>de la media | F-Valor | Pr > F |
|-----------------|---------------|-------------------------|---------|--------|
| Model           | 1             | 2923.92924              | 117.02  | <.0001 |
| Error           | 510           | 12744                   |         |        |
| Total corregido | 511           | 15667                   |         |        |

| Variable  | Estimador del<br>parámetro | Error<br>estándar | Tipo II SS | F-Valor | Pr > F |
|-----------|----------------------------|-------------------|------------|---------|--------|
| Intercept | 3.49314                    | 0.39852           | 1919.77188 | 76.83   | <.0001 |
| semt      | 3.30784                    | 0.30579           | 2923.92924 | 117.02  | <.0001 |

Límites en el número de la condición: 1, 1

Selección escalonada: Paso 2

Variable p19mcos introducida: R-cuadrado = 0.1997 y C(p) = 13.2369

Análisis de la varianza

| Fuente          | Suma de<br>DF | Cuadrado<br>de la media | F-Valor | Pr > F |
|-----------------|---------------|-------------------------|---------|--------|
| Model           | 2             | 3128.14422              | 63.49   | <.0001 |
| Error           | 509           | 12539                   |         |        |
| Total corregido | 511           | 15667                   |         |        |

| Variable  | Estimador del<br>parámetro | Error<br>estándar | Tipo II SS | F-Valor | Pr > F |
|-----------|----------------------------|-------------------|------------|---------|--------|
| Intercept | 3.05097                    | 0.42446           | 1272.79874 | 51.67   | <.0001 |
| semt      | 3.27907                    | 0.30379           | 2870.17748 | 116.51  | <.0001 |
| p19mcos   | 0.10107                    | 0.03510           | 204.21498  | 8.29    | 0.0042 |

Límites en el número de la condición: 1.0011, 4.0043

Selección escalonada: Paso 3

Variable p18msem introducida: R-cuadrado = 0.2132 y C(p) = 6.4277

Análisis de la varianza

| Fuente          | Suma de<br>DF | Cuadrado<br>de la media | F-Valor | Pr > F |
|-----------------|---------------|-------------------------|---------|--------|
| Model           | 3             | 3340.88221              | 45.89   | <.0001 |
| Error           | 508           | 12327                   |         |        |
| Total corregido | 511           | 15667                   |         |        |

| Variable  | Estimador del<br>parámetro | Error<br>estándar | Tipo II SS | F-Valor | Pr > F |
|-----------|----------------------------|-------------------|------------|---------|--------|
| Intercept | 4.41288                    | 0.62371           | 1214.66143 | 50.06   | <.0001 |
| semt      | 3.26031                    | 0.30157           | 2836.17002 | 116.88  | <.0001 |
| p18msem   | -0.16616                   | 0.05612           | 212.73798  | 8.77    | 0.0032 |
| p19mcos   | 0.15991                    | 0.04011           | 385.71984  | 15.90   | <.0001 |

Límites en el número de la condición: 1.3268, 10.961

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1000 significance level for entry into the model.

Resumen de Selección escalonada

| Paso | Variable<br>introducida | Variable<br>eliminada | Número<br>vars en | R-cuadrado<br>parcial | R cuadrado<br>del modelo | C(p)    | F-Valor | Pr > F |
|------|-------------------------|-----------------------|-------------------|-----------------------|--------------------------|---------|---------|--------|
| 1    | semt                    |                       | 1                 | 0.1866                | 0.1866                   | 19.6932 | 117.02  | <.0001 |
| 2    | p19mcos                 |                       | 2                 | 0.0130                | 0.1997                   | 13.2369 | 8.29    | 0.0042 |
| 3    | p18msem                 |                       | 3                 | 0.0136                | 0.2132                   | 6.4277  | 8.77    | 0.0032 |



## Comentario

Los resultados de la Tabla 26, presenta la selección de variables a P: 0.10, de ella se visualiza en tres etapas, en la primera etapa fue testada la variable semilla empleada (semt), en la segunda etapa se incluyó mes de cosecha (p19mcos) y en la tercera etapa mes de siembra (p18mse), ellas poseen significación a P: 0.001.

## Supuestos del análisis de regresión

En el archivo “**02 Base cobertura PAR II jueves.sas**” contiene información de EMPODERAR relacionado al número de beneficiario disgregada por sexo, origen e idioma; también incluye información de presupuestos financiero y en hitos transferidos (Dirección General de Planificación, 2018). En la Tabla 27, se propone el procedimiento REG con la variable dependiente hitopar (presupuesto financiero) con las covariables número de beneficiarios mujeres y condición de origen e idioma; en opciones se solicita límites de confianza, tolerancia (inverso del factor inflacionario de la varianza), factor inflacionario de la varianza, colinealidad, coeficiente de correlación parcial al cuadrado y heteroscedasticidad; además generar un nuevo data sas con los residuos. Por último, con el procedimiento INIVARIATE los residuos serán testados si tiene distribución normal. Los resultados se exponen en la Tabla 28.

Tabla 27. Procedimiento y sentencia de REG para datos de presupuesto financiero en función al número de beneficiarios mujeres y condición de origen e idioma. Además incluye prueba de supuestos del análisis de regresión.

```
proc reg data=par2;
model hitopar= bmuoj omigr onativ castl quech aymar guaran / clb tol vif
collin dwprob r PCORR1 PCORR2 PARTIALR2 HCC;
output out=res r=residus;
run;
quit;

proc univariate normal data=res;
var residus;
run;
```

Tabla 28. Resultados del procedimiento y sentencia de REG para datos de presupuesto financiero en función al número de beneficiarios mujeres y condición de origen e idioma. Además incluye prueba de supuestos del análisis de regresión.

|                               |    |
|-------------------------------|----|
| Base cobertura PAR II         | 11 |
| Procedimiento REG             |    |
| Modelo: MODEL1                |    |
| Variable dependiente: hitopar |    |



**Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



|                                                |                                   |                                    |                                     |                            |                     |                         |
|------------------------------------------------|-----------------------------------|------------------------------------|-------------------------------------|----------------------------|---------------------|-------------------------|
| Número de observaciones leídas                 |                                   | 567                                |                                     |                            |                     |                         |
| Número de observaciones usadas                 |                                   | 567                                |                                     |                            |                     |                         |
| Análisis de la varianza                        |                                   |                                    |                                     |                            |                     |                         |
| Fuente                                         | Suma de<br>DF                     | Cuadrado<br>de la media            | F-Valor                             | Pr > F                     |                     |                         |
| Model                                          | 7                                 | 1.628179E13                        | 2.325971E12                         | 245.10                     | <.0001              |                         |
| Error                                          | 559                               | 5.304847E12                        | 9489886447                          |                            |                     |                         |
| Total corregido                                | 566                               | 2.158664E13                        |                                     |                            |                     |                         |
| Raíz MSE                                       |                                   | 97416                              | R-cuadrado                          | 0.7543                     |                     |                         |
| Media dependiente                              |                                   | 360875                             | Adj R-Sq                            | 0.7512                     |                     |                         |
| Var Coef                                       |                                   | 26.99441                           |                                     |                            |                     |                         |
| Estimadores de parámetros                      |                                   |                                    |                                     |                            |                     |                         |
| -Constante de heteroscedasticidad-             |                                   |                                    |                                     |                            |                     |                         |
| Variable                                       | Estimador del<br>DF               | Error<br>estándar                  | Error<br>Valor t                    | Pr >  t                    | estándar<br>Valor t | Pr >  t                 |
| Intercept                                      | 1                                 | 25869                              | 9570.03166                          | 2.70                       | 0.0071              | 17222 1.50 0.1336       |
| bmuj                                           | 1                                 | 4197.26799                         | 549.69671                           | 7.64                       | <.0001              | 1329.90189 3.16 0.0017  |
| omigr                                          | 1                                 | 8262.53147                         | 588.38381                           | 14.04                      | <.0001              | 798.32109 10.35 <.0001  |
| onativ                                         | 1                                 | 8275.33791                         | 450.66403                           | 18.36                      | <.0001              | 908.97731 9.10 <.0001   |
| castl                                          | 1                                 | 989.15783                          | 356.02217                           | 2.78                       | 0.0056              | 433.42552 2.28 0.0229   |
| quech                                          | 1                                 | -189.54337                         | 406.21857                           | -0.47                      | 0.6410              | 574.29755 -0.33 0.7415  |
| aymar                                          | 1                                 | -1542.17728                        | 381.94348                           | -4.04                      | <.0001              | 809.83876 -1.90 0.0574  |
| guaran                                         | 1                                 | -2182.38046                        | 703.84282                           | -3.10                      | 0.0020              | 1610.15223 -1.36 0.1758 |
| Estimadores de parámetros                      |                                   |                                    |                                     |                            |                     |                         |
| Variable                                       | Corr tipo I<br>semi-parcial<br>DF | Corr tipo I<br>parcial<br>cuadrado | Corr tipo II<br>parcial<br>cuadrada | Inflación de<br>Tolerancia | la varianza         |                         |
| Intercept                                      | 1                                 | .                                  | .                                   | 0                          |                     |                         |
| bmuj                                           | 1                                 | 0.38244                            | 0.38244                             | 0.09445                    | 0.61719             | 1.62024                 |
| omigr                                          | 1                                 | 0.01035                            | 0.01677                             | 0.26078                    | 0.35324             | 2.83093                 |
| onativ                                         | 1                                 | 0.34189                            | 0.56305                             | 0.37624                    | 0.23336             | 4.28516                 |
| castl                                          | 1                                 | 0.00705                            | 0.02658                             | 0.01362                    | 0.39503             | 2.53148                 |
| quech                                          | 1                                 | 0.00294                            | 0.01140                             | 0.00038933                 | 0.30944             | 3.23163                 |
| aymar                                          | 1                                 | 0.00535                            | 0.02094                             | 0.02834                    | 0.71217             | 1.40417                 |
| guaran                                         | 1                                 | 0.00423                            | 0.01691                             | 0.01691                    | 0.71994             | 1.38900                 |
| Estimadores de parámetros                      |                                   |                                    |                                     |                            |                     |                         |
| Constante de<br>heteroscedasticidad 95%        |                                   |                                    |                                     |                            |                     |                         |
| Variable                                       | DF                                | 95% Límites de confianza           | Límites de confianza                |                            |                     |                         |
| Intercept                                      | 1                                 | 7071.34009                         | 44667                               | -7958.86957                | 59697               |                         |
| bmuj                                           | 1                                 | 3117.54447                         | 5276.99151                          | 1585.05235                 | 6809.48363          |                         |
| omigr                                          | 1                                 | 7106.81811                         | 9418.24483                          | 6694.45577                 | 9830.60717          |                         |
| onativ                                         | 1                                 | 7390.13606                         | 9160.53977                          | 6489.90940                 | 10061               |                         |
| castl                                          | 1                                 | 289.85309                          | 1688.46257                          | 137.81614                  | 1840.49951          |                         |
| quech                                          | 1                                 | -987.44471                         | 608.35798                           | -1317.58826                | 938.50152           |                         |
| aymar                                          | 1                                 | -2292.39708                        | -791.95748                          | -3132.87618                | 48.52162            |                         |
| guaran                                         | 1                                 | -3564.88035                        | -799.88057                          | -5345.06853                | 980.30760           |                         |
| Diagnósticos de colinealidad                   |                                   |                                    |                                     |                            |                     |                         |
| Índice de -----Proporción de la variación----- |                                   |                                    |                                     |                            |                     |                         |
| Número                                         | Autovalor                         | la condición                       | Intercept                           | bmuj                       | omigr               | onativ                  |
| 1                                              | 4.30979                           | 1.00000                            | 0.00859                             | 0.01044                    | 0.00520             | 0.00344                 |
| 2                                              | 1.15055                           | 1.93542                            | 0.00004380                          | 0.00850                    | 0.02865             | 0.00552                 |
| 3                                              | 1.00245                           | 2.07347                            | 0.00026728                          | 0.00001335                 | 0.00179             | 0.00003772              |
| 4                                              | 0.74715                           | 2.40173                            | 0.00046206                          | 0.00077433                 | 0.14451             | 0.01583                 |
| 5                                              | 0.40408                           | 3.26584                            | 0.00732                             | 0.00250                    | 0.06971             | 0.00488                 |
| 6                                              | 0.20488                           | 4.58651                            | 0.15566                             | 0.84914                    | 0.00606             | 0.01182                 |
| 7                                              | 0.13636                           | 5.62189                            | 0.79884                             | 0.07786                    | 0.02994             | 0.03897                 |
| 8                                              | 0.04474                           | 9.81435                            | 0.02882                             | 0.05077                    | 0.71414             | 0.91950                 |
| Diagnósticos de colinealidad                   |                                   |                                    |                                     |                            |                     |                         |
| -----Proporción de la variación-----           |                                   |                                    |                                     |                            |                     |                         |
| Número                                         | castl                             | quech                              | aymar                               | guaran                     |                     |                         |
| 1                                              | 0.00583                           | 0.00469                            | 0.00870                             | 0.00154                    |                     |                         |
| 2                                              | 0.00378                           | 0.01051                            | 0.15275                             | 0.20816                    |                     |                         |
| 3                                              | 0.00432                           | 0.07047                            | 0.10064                             | 0.30165                    |                     |                         |
| 4                                              | 0.02753                           | 0.02075                            | 0.00002468                          | 0.16662                    |                     |                         |
| 5                                              | 0.10583                           | 0.06636                            | 0.59128                             | 0.00004511                 |                     |                         |
| 6                                              | 0.00607                           | 0.03930                            | 0.00579                             | 0.08791                    |                     |                         |
| 7                                              | 0.21218                           | 0.11839                            | 0.01982                             | 0.03941                    |                     |                         |
| 8                                              | 0.63446                           | 0.66953                            | 0.12099                             | 0.19467                    |                     |                         |
| Base cobertura PAR II                          |                                   |                                    |                                     | 12                         |                     |                         |
| Procedimiento REG                              |                                   |                                    |                                     |                            |                     |                         |



**Fortalecimiento de la capacidad institucional en los sectores de  
desarrollo integral con coca, tráfico ilícito de drogas y  
seguridad alimentaria para una eficiente gestión del apoyo  
presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



Modelo: MODEL1

Variable dependiente: hitopar

D de Durbin-Watson1.149

Pr < DW<.0001

Pr > DW1.0000

Número de observaciones567

Autocorrelación de orden0.426

NOTA: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

Base cobertura PAR II13

Procedimiento REG

Modelo: MODEL1

Variable dependiente: hitopar

Estadísticos de salida

| Observación | Variable dependiente | Valor  | Error std predicho | Residual Media predicha | Residual | Residual de error std | de Student |
|-------------|----------------------|--------|--------------------|-------------------------|----------|-----------------------|------------|
| 1           | 432340               | 408440 | 13837              | 23900                   | 96428    | 0.248                 |            |
| 2           | 366567               | 361655 | 8164               | 4912                    | 97073    | 0.0506                |            |
| 3           | 236562               | 305546 | 10574              | -68984                  | 96841    | -0.712                |            |
| 4           | 240082               | 274446 | 6525               | -34364                  | 97197    | -0.354                |            |
| 5           | 289580               | 276275 | 5611               | 13305                   | 97254    | 0.137                 |            |
| 6           | 260845               | 261728 | 5324               | -883.3666               | 97270    | -0.0091               |            |
| .....       |                      |        |                    |                         |          |                       |            |
| .....       |                      |        |                    |                         |          |                       |            |

Estadísticos de salida

| Observación | D     | de Cook |
|-------------|-------|---------|
| 1           | 0.000 |         |
| 2           | 0.000 |         |
| 3           | 0.001 |         |
| 4           | 0.000 |         |
| 5           | 0.000 |         |
| .....       |       |         |
| .....       |       |         |

Suma de residuales0

Suma de los residuales cuadrados5.304847E12

SS de Residual predicho (PRESS)6.122001E12

Base cobertura PAR II17

Procedimiento UNIVARIATE

Variable: residus (Residual)

Momentos

|                 |            |                 |            |
|-----------------|------------|-----------------|------------|
| N               | 567        | Sumar pesos     | 567        |
| Media           | 0          | Observ suma     | 0          |
| Desviación std  | 96811.778  | Varianza        | 9372520361 |
| Asimetría       | -2.2723032 | Curtosis        | 10.4476887 |
| SC no corregida | 5.30485E12 | SC corregida    | 5.30485E12 |
| Coef. variación | .          | Media error std | 4065.71252 |

Medidas estadísticas básicas

| Ubicación | Varianza           |
|-----------|--------------------|
| Media     | 0.00               |
| Mediana   | 20007.55           |
| Moda      | .                  |
|           | Rango              |
|           | Rango intercuantil |

Tests para posición: Mu0=0

| Test                 | -Estadístico- | -----P-valor----- |
|----------------------|---------------|-------------------|
| T de Student         | t             | 0                 |
| Signo                | M             | 84.5              |
| Puntuación con signo | S             | 16734             |

Tests para normalidad

| Test               | -Estadístico-- | -----P-valor----- |
|--------------------|----------------|-------------------|
| Shapiro-Wilk       | W              | 0.814156          |
| Kolmogorov-Smirnov | D              | 0.154083          |
| Cramer-von Mises   | W-Sq           | 4.846583          |
| Anderson-Darling   | A-Sq           | 26.00345          |



| Cuantiles (Definición 5) |             |                    |             |
|--------------------------|-------------|--------------------|-------------|
| Cuantil                  |             | Estimador          |             |
| 100% Máx                 |             | 346933.5           |             |
| 99%                      |             | 173672.2           |             |
| 95%                      |             | 106923.8           |             |
| 90%                      |             | 85412.1            |             |
| 75% Q3                   |             | 48779.1            |             |
| 50% Mediana              |             | 20007.6            |             |
| 25% Q1                   |             | -27373.1           |             |
| 10%                      |             | -92871.2           |             |
| 5%                       |             | -172569.0          |             |
| 1%                       |             | -324558.6          |             |
| 0% Mín                   |             | -623970.1          |             |
|                          |             |                    |             |
| Observaciones extremas   |             |                    |             |
| -----Inferior-----       |             | -----Superior----- |             |
| Valor                    | Observación | Valor              | Observación |
| -623970                  | 101         | 176466             | 250         |
| -616979                  | 553         | 178661             | 183         |
| -603178                  | 458         | 199687             | 252         |
| -459504                  | 554         | 284127             | 275         |
| -426292                  | 106         | 346933             | 545         |

## Comentario

Los resultados (Tabla 28), considera 567 observaciones; en el análisis de varianza el modelo debido a la regresión es significativa a  $P: 0.01$ ; el coeficiente de determinación  $R^2$  presenta un 75.43 % es decir las variación de las covariables explican en dicho porcentaje sobre la variación de hito presupuestario (variable dependiente); los coeficientes de regresión a nivel poblacional diferente de cero a  $P: 0.01$ , excepto los beneficiarios que hablan quechua; según la prueba de heteroscedasticidad son significativas excepto beneficiarios que hablan quechua, aymara y guaraní. Los coeficientes de correlación parcial cuadrada presentan valores mayores con beneficiarios mujeres y origen migrante. Referente a la tolerancia las próximas a la unidad son beneficiarios mujeres, hablan aymara y guaraní deduciéndose que tiene escasa colinealidad. Los límites de confianza tanto de los coeficientes de regresión a nivel poblacional reflejan que son diferentes de cero. El coeficiente de D de Durbin-Watson cita una autocorrelación positiva (significativa a  $P: 0.01$ ). También se generó los valores predicho, residual y el coeficiente de D de Cook.

El procedimiento UNIVARIATE presenta una media 0 y la prueba de normalidad de Shapiro-Wilk es significativa, es decir los residuos presentan distribución diferente a la normal a  $P: 0.01$ .

## 4.2. Análisis de correspondencia

El análisis de correspondencia simple y un análisis de correspondencia múltiple pueden usarse para encontrar una representación gráfica de baja dimensión de las filas y columnas de una tabla de contingencia. Posee el objetivo de registrar las relaciones entre dos conjuntos de variables.

Como ejemplo, vemos los datos del archivo “**Datos de papa evaluado 2017-2018 jueves 29.sas**” de la encuesta de papa en Cochabamba (OAP-MDRyT, 2018), empleándose las variables variedad y P16 ¿Cómo califica la cosecha del año agrícola? ambas de la campaña 2017-2018. En el archivo (Tabla 29) se incluyó tres nuevas variables con observaciones literales referidas a variedad (varie), calificación de la cosecha (p16cal) y aplicación de riego (riego). En la Tabla 30, con el procedimiento CORRESP se pide



el análisis de correspondencia simple de las variables variedad (varie) y calificación de la cosecha (p16cal), así mismo los valores observados y esperados se generan en un archivo sas. Considerando los datos de las dimensiones 1 y 2 con el procedimiento GPLOT se gráfica y se imprime con PRINT. Los resultados se expone en la Tabla 31.

Tabla 29. Datos en archivo sas de la Encuesta Datos de papa evaluado 2017-2018 en Cochabamba (OAP-MDRyT, 2018).

```
data papaen;
input produce 1-3 p01dep 6 p02pro 9-11 p03mun 13-17 p05prod $ 20-51 folio 58-60 varied sup1718
rend1718 p16ccalf semt p18mse p19mcos p20fadv p20plg1 p20enf1 p23agrqm p24guan p24barr
p24zanj p24aver p24comp p24rot p24labr p24terr p24otr p25cods p26dest p27apoy p30rieg ;

if p16ccalf=1 then p16cal='Bueno'; if p16ccalf=2 then p16cal='Regular'; if p16ccalf=3 then p16cal='Malo';
if varied=1 then varie='Waycha'; if varied=2 then varie='Imilla'; if varied=3 then varie='Sani';
if varied=4 then varie='Desire'; if varied=5 then varie='Nativa'; if varied=6 then varie='Otros';
if p30rieg=1 then riego='Si'; if p30rieg=2 then riego='No';

title1 'Encuesta papa Cochabamba (OAP-MDRyT, 2018)';
cards;
1 7 708 70801 ISIDRO ROJAS 169 4 12.00 4.79 2 0.38 10 2 1 98 98 1 2 2 2 1 2 2 2 2 2 2 2 1
1 2 213 21301 SILVANO MORALES ALBERTO 129 1 10.00 5.00 2 0.28 10 4 1 98 98 1 1 2 2 1 2 2 2 2 2 1 2 2
1 7 705 70803 JUAN CARLOS ROJAS 168 4 8.00 7.19 2 1.44 11 2 1 98 98 1 2 2 2 1 2 2 2 2 2 2 2 2
.....
```

Tabla 30. Procedimiento y sentencia de CORRESP para la variable variedad y la percepción ¿P16. Cómo califica la cosecha del año agrícola? ambas de la campaña 2017-2018.

```
proc corresp data=papaen out=paco short observed expected ;
tables varie, p16cal;
run;

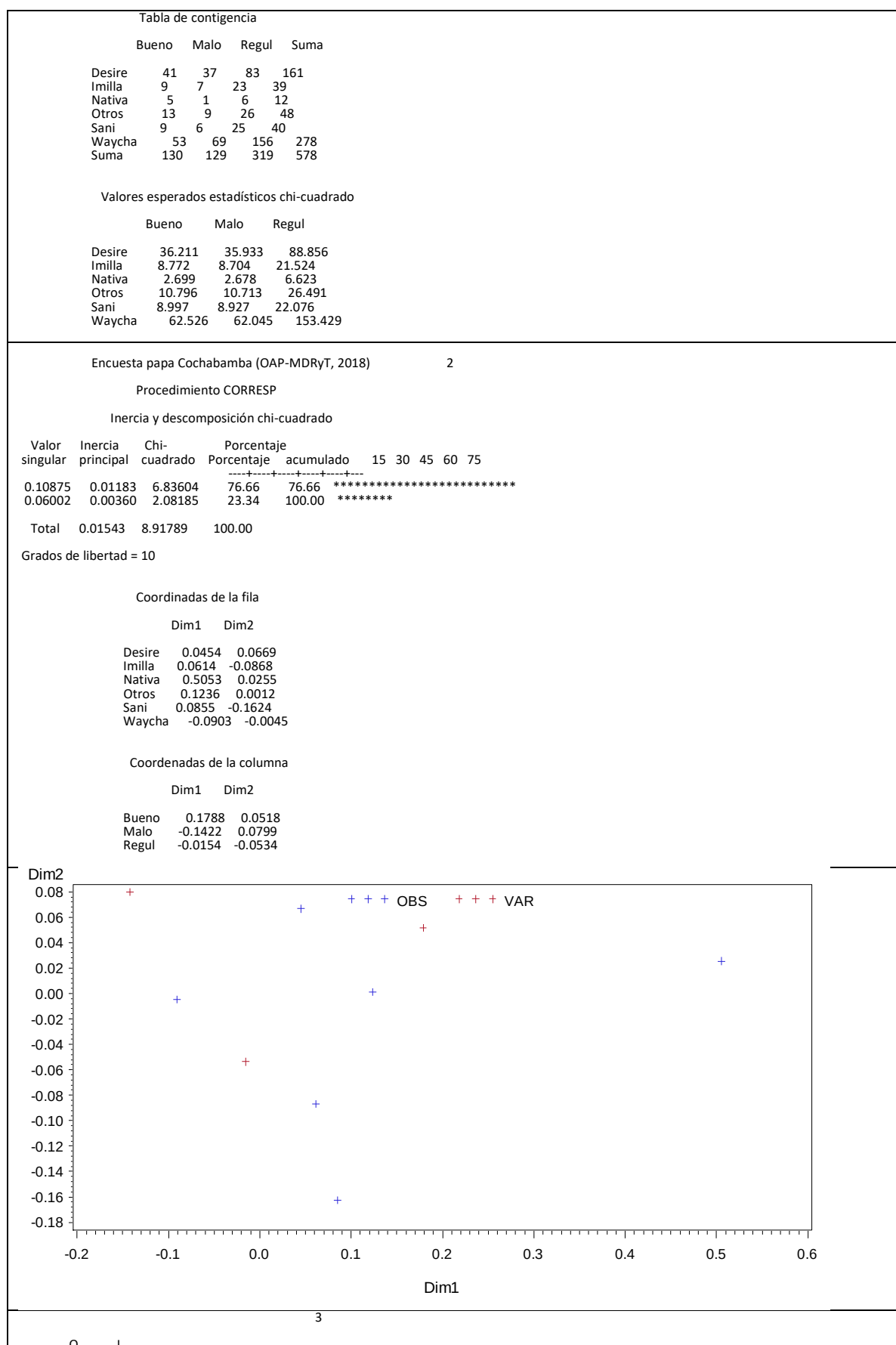
ods graphics off;
goptions reset=all border;
legend1 label= none
position=(top center inside)
mode=share;

data pa; set paco; if _type_='INERTIA' then delete; run;

proc gplot data=pa ;
plot dim2*dim1=_type_ / legend=legend1 ;
run;
proc print data=pa;
run;
```

Tabla 31. Salida del procedimiento y sentencia de CORRESP para la variable variedad y la percepción ¿P16. Cómo califica la cosecha del año agrícola? ambas de la campaña 2017-2018.

|                                            |   |
|--------------------------------------------|---|
| Encuesta papa Cochabamba (OAP-MDRyT, 2018) | 1 |
| Procedimiento CORRESP                      |   |





[illegible]

## Comentario

Los resultados de la Tabla 31, presenta la distribución de frecuencias observadas; frecuencias esperadas; la variación explicada del primer componente en un 76.66%; las coordenadas de las filas y columnas; la gráfica de dispersión donde se los puntos no están próximas entre ellas, por lo cual se deduce que no existe relación entre las clases de ambas variables.

## Análisis de correspondencia múltiple

En la Tabla 32, el procedimiento CORRESP presenta sentencias para el análisis de correspondencia múltiple incluyendo las variables variedad, la percepción ¿P16. Cómo califica la cosecha del año agrícola de la campaña 2017-2018? y si aplico riego?.

Tabla 32. Procedimiento y sentencia de CORRESP para la variable variedad, la percepción ¿P16. Cómo califica la cosecha del año agrícola de la campaña 2017-2018? y si aplico riego?.

```
proc corresp mca data=papaen out=pac short OBSERVED expected ;
    tables varie p16cal riego;
run;

data pam; set pac; if _type_='INERTIA' then delete; run;
proc gplot data=pam ;
plot dim2*dim1=_type_ / legend=legend1 ;
run;
proc print data=pam;
run;
```

Tabla 33. Salida del procedimiento y sentencia de CORRESP para variedad, la percepción ¿P16. Cómo califica la cosecha del año agrícola? y riego de la campaña 2017-2018.

|                       |        |        |       |      |        |       |      |       |    |    |
|-----------------------|--------|--------|-------|------|--------|-------|------|-------|----|----|
| Procedimiento CORRESP |        |        |       |      |        |       |      |       |    |    |
| Tabla Burt            |        |        |       |      |        |       |      |       |    |    |
| Desire                | Imilla | Nativa | Otros | Sani | Waycha | Buono | Male | Regul | No | Si |



**Fortalecimiento de la capacidad institucional en los sectores de  
desarrollo integral con coca, tráfico ilícito de drogas y  
seguridad alimentaria para una eficiente gestión del apoyo  
presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



|        |     |    |    |    |    |     |     |     |     |     |     |
|--------|-----|----|----|----|----|-----|-----|-----|-----|-----|-----|
| Desire | 160 | 0  | 0  | 0  | 0  | 0   | 41  | 36  | 83  | 87  | 73  |
| Imilla | 0   | 39 | 0  | 0  | 0  | 0   | 9   | 7   | 23  | 31  | 8   |
| Nativa | 0   | 0  | 12 | 0  | 0  | 0   | 5   | 1   | 6   | 10  | 2   |
| Otros  | 0   | 0  | 0  | 48 | 0  | 0   | 13  | 9   | 26  | 36  | 12  |
| Sani   | 0   | 0  | 0  | 0  | 40 | 0   | 9   | 6   | 25  | 31  | 9   |
| Waycha | 0   | 0  | 0  | 0  | 0  | 278 | 53  | 69  | 156 | 185 | 93  |
| Bueno  | 41  | 9  | 5  | 13 | 9  | 53  | 130 | 0   | 0   | 90  | 40  |
| Malo   | 36  | 7  | 1  | 9  | 6  | 69  | 0   | 128 | 0   | 80  | 48  |
| Regul  | 83  | 23 | 6  | 26 | 25 | 156 | 0   | 0   | 319 | 210 | 109 |
| No     | 87  | 31 | 10 | 36 | 31 | 185 | 90  | 80  | 210 | 380 | 0   |
| Si     | 73  | 8  | 2  | 12 | 9  | 93  | 40  | 48  | 109 | 0   | 197 |

Valores esperados estadísticos chi-cuadrado

|        | Desire  | Imilla | Nativa | Otros  | Sani   | Waycha  |
|--------|---------|--------|--------|--------|--------|---------|
| Desire | 44.367  | 10.815 | 3.328  | 13.310 | 11.092 | 77.088  |
| Imilla | 10.815  | 2.636  | 0.811  | 3.244  | 2.704  | 18.790  |
| Nativa | 3.328   | 0.811  | 0.250  | 0.998  | 0.832  | 5.782   |
| Otros  | 13.310  | 3.244  | 0.998  | 3.993  | 3.328  | 23.127  |
| Sani   | 11.092  | 2.704  | 0.832  | 3.328  | 2.773  | 19.272  |
| Waycha | 77.088  | 18.790 | 5.782  | 23.127 | 19.272 | 133.941 |
| Bueno  | 36.049  | 8.787  | 2.704  | 10.815 | 9.012  | 62.634  |
| Malo   | 35.494  | 8.652  | 2.662  | 10.648 | 8.873  | 61.671  |
| Regul  | 88.458  | 21.562 | 6.634  | 26.537 | 22.114 | 153.695 |
| No     | 105.373 | 25.685 | 7.903  | 31.612 | 26.343 | 183.085 |
| Si     | 54.627  | 13.315 | 4.097  | 16.388 | 13.657 | 94.915  |

Valores esperados estadísticos chi-cuadrado

|        | Bueno  | Malo   | Regul   | No      | Si      |
|--------|--------|--------|---------|---------|---------|
| Desire | 36.049 | 35.494 | 88.458  | 105.373 | 54.627  |
| Imilla | 8.787  | 8.652  | 21.562  | 25.685  | 13.315  |
| Nativa | 2.704  | 2.662  | 6.634   | 7.903   | 4.097   |
| Otros  | 10.815 | 10.648 | 26.537  | 31.612  | 16.388  |
| Sani   | 9.012  | 8.873  | 22.114  | 26.343  | 13.657  |
| Waycha | 62.634 | 61.671 | 153.695 | 183.085 | 94.915  |
| Bueno  | 29.289 | 28.839 | 71.872  | 85.615  | 44.385  |
| Malo   | 28.839 | 28.395 | 70.766  | 84.298  | 43.702  |
| Regul  | 71.872 | 70.766 | 176.362 | 210.087 | 108.913 |
| No     | 85.615 | 84.298 | 210.087 | 250.260 | 129.740 |
| Si     | 44.385 | 43.702 | 108.913 | 129.740 | 67.260  |

Procedimiento CORRESP

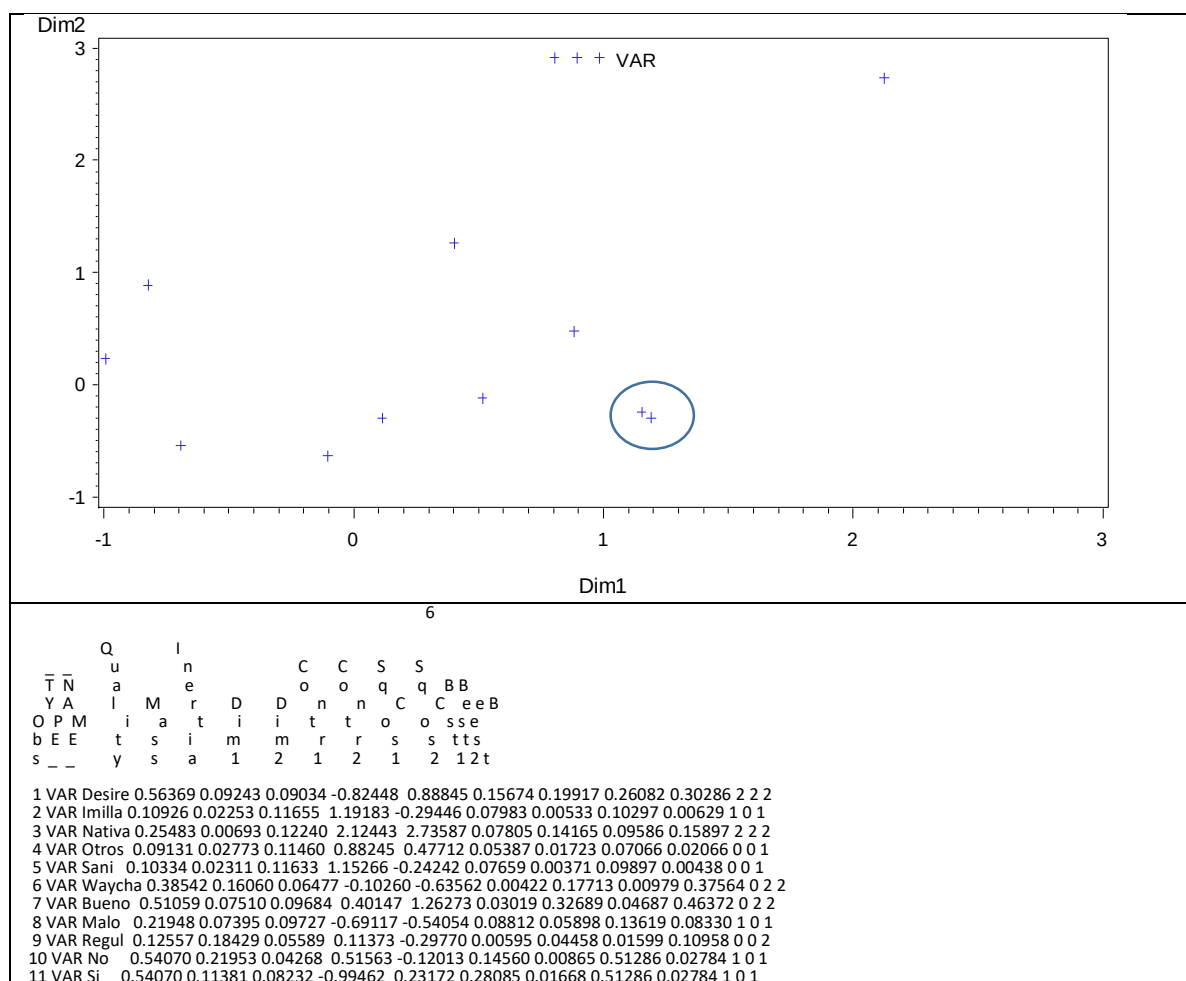
Inercia y descomposición chi-cuadrado

| Valor singular | Inercia principal | Chi-cuadrado | Porcentaje Porcentaje | acumulado | 3     | 6 | 9 | 12 | 15 |
|----------------|-------------------|--------------|-----------------------|-----------|-------|---|---|----|----|
| 0.63315        | 0.40088           | 702.54       | 15.03                 | 15.03     | ***** |   |   |    |    |
| 0.60525        | 0.36632           | 641.98       | 13.74                 | 28.77     | ***** |   |   |    |    |
| 0.58554        | 0.34286           | 600.87       | 12.86                 | 41.63     | ***** |   |   |    |    |
| 0.57735        | 0.33333           | 584.17       | 12.50                 | 54.13     | ***** |   |   |    |    |
| 0.57735        | 0.33333           | 584.17       | 12.50                 | 66.63     | ***** |   |   |    |    |
| 0.56725        | 0.32177           | 563.91       | 12.07                 | 78.69     | ***** |   |   |    |    |
| 0.54473        | 0.29673           | 520.03       | 11.13                 | 89.82     | ***** |   |   |    |    |
| 0.52099        | 0.27144           | 475.69       | 10.18                 | 100.00    | ***** |   |   |    |    |
| Total          | 2.66667           | 4673.36      | 100.00                |           |       |   |   |    |    |

Grados de libertad = 100

Coordenadas de la columna

|        | Dim1    | Dim2    |
|--------|---------|---------|
| Desire | -0.8245 | 0.8884  |
| Imilla | 1.1918  | -0.2945 |
| Nativa | 2.1244  | 2.7359  |
| Otros  | 0.8824  | 0.4771  |
| Sani   | 1.1527  | -0.2424 |
| Waycha | -0.1026 | -0.6356 |
| Bueno  | 0.4015  | 1.2627  |
| Malo   | -0.6912 | -0.5405 |
| Regul  | 0.1137  | -0.2977 |
| No     | 0.5156  | -0.1201 |
| Si     | -0.9946 | 0.2317  |



En la Tabla 33, los resultados reflejan la distribución de frecuencias observadas, frecuencias esperadas; el primer componente explica la variación en un 15.03 % y los primeros cuatro componentes explican la variación en un 54.13 %. En función a las coordenadas en la figura los puntos dentro del círculo están ciertamente relacionadas y corresponden a las variedades de papa Sani e Imilla. Posteriormente se expone los valores de las dimensiones y otros estadísticos.

### 4.3. Análisis de agrupación

El análisis de cluster tiene una serie de técnicas multivariadas, que tienen el objetivo principal de agrupar objetos en función a las características que poseen. En el programa sas el procedimiento CLUSTER presenta los métodos AVERAGE, CENTROID, COMPLETE, DENSITY, ELM, FLEXIBLE, MCQUITTY, MEDIAN, SINGLE, TWOSTAGE y WARD.

### Cluster Método Wards (Variables cuantitativas)

En el método de varianza mínima de Ward, la distancia entre dos grupos es la suma de cuadrados ANOVA entre los dos grupos sumada sobre todas las variables. En cada generación, la suma de



cuadrados dentro del grupo se minimiza sobre todas las particiones que se pueden obtener al fusionar dos grupos de la generación anterior. Las sumas de cuadrados son más fáciles de interpretar cuando se dividen por la suma total de cuadrados para dar proporciones de varianza (correlaciones semiparciales al cuadrado). El método de Ward se une a grupos para maximizar la probabilidad en cada nivel de la jerarquía bajo los siguientes supuestos: mezcla normal multivariable; matrices de covarianza esféricas iguales; y probabilidades de muestreo iguales. El método de Ward tiende a unir grupos con un pequeño número de observaciones, y está fuertemente sesgado hacia la producción de grupos con aproximadamente el mismo número de observaciones. También es muy sensible a los valores atípicos (Milligan 1980).

El Instituto Nacional de Estadística de Bolivia, expone los datos de la Encuesta de Hogares 2017, de ella se extrajo información referida a Gastos no alimentarios (Sección 10: gastos; Parte B: Gastos del hogar) la misma se expone en el archivo **EH2017\_GastosNoAlimentarios.xlsx** (incluye la planilla con las interrogantes) y archivo **“gastonoalimen.sas”**. En el SAS, luego de abrir el archivo gastonoalimen.sas (Tabla 34); en la Tabla 35 se expone el procedimiento CLUSTER en ella se cita el método de agrupación Varianza Mínima de Wards, genera un nuevo data sas para ser empleada por el procedimiento TREE, la misma graficará el Dendrograma y un nuevo data sas con la asignación de los grupos; y en la Tabla 36 se expone los resultados.

Tabla 34. Procedimiento y sentencia de DATA para el grupo de datos del gasto no alimentario dentro el hogar de varios productos (Encuesta de Hogar. INE, 2017).

```
Data gastono;
input folio depto area s10b1001 s10b1002 s10b1003 s10b1004 s10b1005 s10b1006 s10b1007
s10b1008 s10b1009 s10b1010 s10b1011 s10b1012 s10b1013 s10b1014 s10b1015 s10b1016 s10b1017
s10b1018 s10b1019 s10b1020 s10b1021 s10b1022 s10b1023 s10b1024 s10b1025 s10b1026 s10b1027
s10b1028;
cards;
666 8 1 0 0 0 0 0 0 0 50 0 15 0 0 0 30 0 0 0 0 200 0 0 100 0 0 0 350 50 0 0
759 8 1 50 0 0 0 0 0 0 50 70 0 0 0 0 0 20 0 0 150 80 0 0 0 0 0 0 50 300 150 150
881 5 1 0 0 0 0 0 0 100 20 0 10 0 0 0 0 0 250 0 0 50 10 0 30 600 0 0 300 50 0 0
.....
```

Tabla 35. Procedimiento y sentencia de CLUSTER y TREE para el gasto alimentario de varios productos dentro el hogar.

```
proc cluster method=ward data=gastono outtree=tree01;
var s10b1001 s10b1002 s10b1003 s10b1004 s10b1005 s10b1006 s10b1007
s10b1008 s10b1009 s10b1010 s10b1011 s10b1012 s10b1013 s10b1014 s10b1015
s10b1016 s10b1017 s10b1018 s10b1019 s10b1020 s10b1021 s10b1022 s10b1024
s10b1025 s10b1026 s10b1027 s10b1028;
id folio;
run;

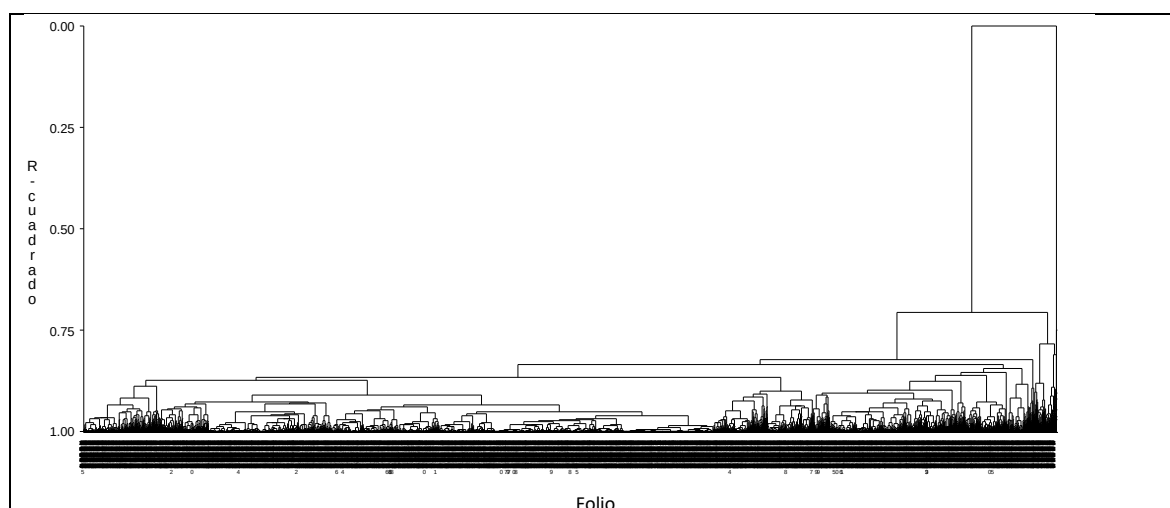
proc tree data=tree01 n=5 out=out01 h=rsq;
id folio;
```



```
run;
quit;
```

Tabla 36. Salidas del procedimiento y sentencia de CLUSTER y TREE para el gasto alimentario de varios productos dentro el hogar.

| Procedimiento CLUSTER                                                       |                              |            |              |        |
|-----------------------------------------------------------------------------|------------------------------|------------|--------------|--------|
| Análisis del conglomerado de la varianza mínima de Ward                     |                              |            |              |        |
| Autovalores de la matriz de covarianza                                      |                              |            |              |        |
| Autovalor                                                                   | Diferencia                   | Proporción | Acumulada    |        |
| 1                                                                           | 1294560.07                   | 1155843.09 | 0.7508       | 0.7508 |
| 2                                                                           | 138716.98                    | 76995.67   | 0.0804       | 0.8312 |
| 3                                                                           | 61721.31                     | 18168.64   | 0.0358       | 0.8670 |
| 4                                                                           | 43552.67                     | 5166.68    | 0.0253       | 0.8923 |
| 5                                                                           | 38385.99                     | 10208.10   | 0.0223       | 0.9145 |
| 6                                                                           | 28177.89                     | 3164.54    | 0.0163       | 0.9309 |
| 7                                                                           | 25013.35                     | 11130.84   | 0.0145       | 0.9454 |
| 8                                                                           | 13882.51                     | 2542.48    | 0.0081       | 0.9534 |
| 9                                                                           | 11340.03                     | 1787.14    | 0.0066       | 0.9600 |
| 10                                                                          | 9552.89                      | 2269.18    | 0.0055       | 0.9655 |
| 11                                                                          | 7283.70                      | 34.85      | 0.0042       | 0.9698 |
| 12                                                                          | 7248.86                      | 755.02     | 0.0042       | 0.9740 |
| 13                                                                          | 6493.84                      | 231.53     | 0.0038       | 0.9777 |
| 14                                                                          | 6262.31                      | 704.36     | 0.0036       | 0.9814 |
| 15                                                                          | 5557.95                      | 855.57     | 0.0032       | 0.9846 |
| 16                                                                          | 4702.39                      | 733.47     | 0.0027       | 0.9873 |
| 17                                                                          | 3968.92                      | 548.00     | 0.0023       | 0.9896 |
| 18                                                                          | 3420.92                      | 396.54     | 0.0020       | 0.9916 |
| 19                                                                          | 3024.39                      | 204.73     | 0.0018       | 0.9934 |
| 20                                                                          | 2819.65                      | 332.56     | 0.0016       | 0.9950 |
| 21                                                                          | 2487.09                      | 836.22     | 0.0014       | 0.9964 |
| 22                                                                          | 1650.87                      | 283.20     | 0.0010       | 0.9974 |
| 23                                                                          | 1367.67                      | 345.90     | 0.0008       | 0.9982 |
| 24                                                                          | 1021.77                      | 183.90     | 0.0006       | 0.9988 |
| 25                                                                          | 837.87                       | 98.53      | 0.0005       | 0.9993 |
| 26                                                                          | 739.33                       | 218.02     | 0.0004       | 0.9997 |
| 27                                                                          | 521.32                       |            | 0.0003       | 1.0000 |
| Desviación estándar muestral total de la raíz cuadrada de la media 252.7121 |                              |            |              |        |
| Distancia de la raíz cuadrada media entre observaciones 1857.047            |                              |            |              |        |
| Historia de conglomerado                                                    |                              |            |              |        |
| T                                                                           |                              |            |              |        |
| i                                                                           |                              |            |              |        |
| NCL                                                                         | ----Conglomerados unidos---- | Frecuencia | SPRSQ        | RSQ e  |
| 11130                                                                       | 4076                         | 6437       | 2 0.0000     | 1.00 T |
| 11129                                                                       | CL11130                      | 6572       | 3 0.0000     | 1.00 T |
| 11128                                                                       | CL11129                      | 6872       | 4 0.0000     | 1.00 T |
| 11127                                                                       | CL11128                      | 7970       | 5 0.0000     | 1.00 T |
| 11126                                                                       | CL11127                      | 8105       | 6 0.0000     | 1.00 T |
| 11125                                                                       | CL11126                      | 8601       | 7 0.0000     | 1.00 T |
| 11124                                                                       | 8941                         | 11203      | 2 0.0000     | 1.00 T |
| .....                                                                       |                              |            |              |        |
| 20                                                                          | CL24                         | CL91       | 1630 0.0023  | .897   |
| 19                                                                          | CL48                         | CL34       | 23 0.0024    | .895   |
| 18                                                                          | CL40                         | CL173      | 55 0.0025    | .893   |
| 17                                                                          | CL28                         | 92128      | 104 0.0026   | .890   |
| 16                                                                          | CL31                         | CL100      | 866 0.0029   | .887   |
| 15                                                                          | CL32                         | CL68       | 254 0.0031   | .884   |
| 14                                                                          | CL22                         | CL37       | 136 0.0034   | .881   |
| 13                                                                          | CL20                         | CL46       | 1742 0.0034  | .877   |
| 12                                                                          | CL16                         | CL27       | 7207 0.0047  | .873   |
| 11                                                                          | CL12                         | CL21       | 8338 0.0051  | .867   |
| 10                                                                          | CL13                         | CL41       | 2216 0.0069  | .861   |
| 9                                                                           | CL10                         | CL492      | 2218 0.0070  | .854   |
| 8                                                                           | CL9                          | CL15       | 2472 0.0086  | .845   |
| 7                                                                           | CL11                         | CL8        | 10810 0.0099 | .835   |
| 6                                                                           | CL7                          | CL18       | 10865 0.0125 | .823   |
| 5                                                                           | CL17                         | CL19       | 127 0.0129   | .810   |
| 4                                                                           | CL14                         | CL5        | 263 0.0249   | .785   |
| 3                                                                           | CL23                         | 87725      | 3 0.0339     | .751   |
| 2                                                                           | CL6                          | CL4        | 11128 0.0450 | .706   |
| 1                                                                           | CL2                          | CL3        | 11131 0.7058 | .000   |



## Comentario

Los resultados de la Tabla 36, en los autovalores presenta una variación explicada en el primer componente principal un 75.08 y 8.04% el segundo componente, entre ambos explican un 83.12 % sobre la variación total. El historial de conglomerado presenta los valores de  $R^2$  semi parcial,  $R^2$ , frecuencia por grupo. La asignación del número de grupos está supeditado a la experiencia del investigador o a los criterios en base a los estadísticos como  $R^2$ . Por ejemplo para agrupar en cinco, se tiene un valor de  $R^2 = 0.81$ . El dendrograma refleja los grupos con mayor detalle y en función al  $R^2$ .

## Componentes principales

El análisis de componentes principales es una técnica multivariable para examinar las relaciones entre varias variables cuantitativas. Debe utilizar el procedimiento PRINCOMP si está interesado en resumir datos y detectar relaciones lineales. Puede usar componentes principales para reducir el número de variables en regresión, agrupación en clústeres, etc.

En la Tabla 37, continuando con los datos del archivo gastonoalimen.sas, el procedimiento PRINCOMP emplea la matriz de covarianza, genera un nuevo data sas; los resultados (nuevo data sas) de TREE y PRINCOMP debe juntarla en otra data sas para generar análisis de las variables considerando los grupos o clusters. Los resultados se presentan en la Tabla 38.

Tabla 37. Procedimiento PRINCOMP y sentencia de Componentes principales para el gasto alimentario de varios productos dentro el hogar y generación de figura.

```
proc princomp data=gastono cov out=prin1 ;
  var s10b1001 s10b1002 s10b1003 s10b1004 s10b1005 s10b1006 s10b1007
  s10b1008 s10b1009 s10b1010 s10b1011 s10b1012 s10b1013 s10b1014 s10b1015
  s10b1016 s10b1017 s10b1018 s10b1019 s10b1020 s10b1021 s10b1022 s10b1024
  s10b1025 s10b1026 s10b1027 s10b1028;
  id folio;
```



run;

```
proc sort data= prin1 ; by folio; run;
proc sort data= out01 ; by folio; run;
data new1; merge prin1 out01 ; by folio; run;
proc print data=new1 ; run;
```

Tabla 38. Salidas del procedimiento y sentencia de componentes principales para el gasto alimentario de varios productos dentro el hogar.

| Procedimiento PRINCOMP |             |             |             |             |             |             |
|------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Observaciones          |             | 11131       |             |             |             |             |
| Variables              |             | 27          |             |             |             |             |
| Estadísticos simples   |             |             |             |             |             |             |
| s10b1001               | s10b1002    | s10b1003    | s10b1004    | s10b1005    | s10b1006    |             |
| Media                  | 46.84969904 | 115.3218040 | 22.36043482 | 19.0486030  | 17.99020753 | 9.53827149  |
| StD                    | 57.87968867 | 165.7287881 | 80.20661209 | 108.6677097 | 89.90340725 | 30.22780340 |
| Estadísticos simples   |             |             |             |             |             |             |
| s10b1007               | s10b1008    | s10b1009    | s10b1010    | s10b1011    | s10b1012    |             |
| Media                  | 45.17231156 | 25.44124517 | 15.88275986 | 31.4556644  | 64.3586830  | 10.94627617 |
| StD                    | 51.30752464 | 60.95604265 | 35.60800611 | 231.4050433 | 364.0043469 | 56.39506063 |
| Estadísticos simples   |             |             |             |             |             |             |
| s10b1013               | s10b1014    | s10b1015    | s10b1016    | s10b1017    | s10b1018    |             |
| Media                  | 55.0747462  | 15.81798581 | 79.0397988  | 4.57088312  | 32.25604169 | 35.79979337 |
| StD                    | 123.4537827 | 62.52900476 | 177.4111613 | 31.29779375 | 96.26527080 | 76.14706186 |
| Estadísticos simples   |             |             |             |             |             |             |
| s10b1019               | s10b1020    | s10b1021    | s10b1022    | s10b1024    | s10b1025    |             |
| Media                  | 8.61741982  | 16.06104573 | 18.86465726 | 23.4297907  | 10.84978888 | 69.8716198  |
| StD                    | 25.41630829 | 80.00848437 | 44.55308458 | 162.9609371 | 86.17648936 | 226.7917840 |
| Estadísticos simples   |             |             |             |             |             |             |
| s10b1026               |             | s10b1027    |             | s10b1028    |             |             |
| Media                  | 84.8424670  | 20.98531129 |             | 48.403647   |             |             |
| StD                    | 110.9068448 | 62.82551376 |             | 1136.682554 |             |             |
| Matriz de covarianza   |             |             |             |             |             |             |
| s10b1001               | s10b1002    | s10b1003    | s10b1004    | s10b1005    | s10b1006    |             |
| s10b1001               | 3350.058    | 1822.042    | 46.536      | 393.520     | 696.044     | 178.984     |
| s10b1002               | 1822.042    | 27466.031   | -242.891    | 2614.845    | 2073.716    | 570.213     |
| s10b1003               | 46.536      | -242.891    | 6433.101    | 84.214      | -91.957     | 44.780      |
| s10b1004               | 393.520     | 2614.845    | 84.214      | 11808.671   | 1202.602    | 219.221     |
| s10b1005               | 696.044     | 2073.716    | -91.957     | 1202.602    | 8082.623    | 502.830     |
| s10b1006               | 178.984     | 570.213     | 44.780      | 219.221     | 502.830     | 913.720     |
| s10b1007               | 1588.929    | 1883.915    | 83.826      | 582.064     | 603.334     | 216.846     |
| s10b1008               | 603.635     | 827.646     | 12.393      | 219.883     | 201.777     | 135.905     |
| s10b1009               | 436.316     | 911.232     | 14.845      | 345.822     | 530.170     | 83.038      |
| s10b1010               | 2214.761    | 1764.019    | -390.020    | 1171.410    | 3157.663    | 60.408      |
| s10b1011               | 2696.608    | 7343.900    | -128.266    | 3955.584    | 4959.766    | 331.760     |
| s10b1012               | 268.513     | 1066.992    | 91.164      | 408.528     | 555.591     | 112.394     |
| s10b1013               | 1047.570    | 3883.831    | 247.976     | 1379.437    | 1401.361    | 261.584     |



**Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



|          |                |              |          |          |           |         |
|----------|----------------|--------------|----------|----------|-----------|---------|
| s10b1014 | 253.229        | 524.048      | -58.800  | 69.803   | 485.760   | 78.942  |
| s10b1015 | 886.544        | 4844.077     | 251.970  | 992.965  | 1905.308  | 284.132 |
| s10b1016 | 28.856         | 335.747      | 6.077    | 35.097   | 135.586   | 27.361  |
| s10b1017 | 349.224        | 893.080      | 154.018  | 230.187  | 919.256   | 160.381 |
| s10b1018 | 881.571        | 2410.144     | 308.918  | 1084.438 | 1027.060  | 261.665 |
| s10b1019 | 247.182        | 576.399      | 48.647   | 218.314  | 365.043   | 61.849  |
| s10b1020 | 352.698        | 680.419      | 218.705  | 118.533  | 673.903   | 263.562 |
| s10b1021 | 298.226        | 1039.637     | 124.487  | 287.875  | 420.783   | 127.202 |
| s10b1022 | 93.487         | -678.344     | 758.826  | 26.420   | 62.233    | 106.494 |
| s10b1024 | 147.196        | 378.837      | 189.501  | 215.260  | 349.599   | 105.244 |
| s10b1025 | 2151.882       | 77.685       | 185.877  | 1107.271 | 1727.878  | 273.033 |
| s10b1026 | 1708.760       | 5606.066     | 510.350  | 1143.609 | 1808.304  | 524.809 |
| s10b1027 | 660.484        | 2029.036     | 22.333   | 468.208  | 921.088   | 95.240  |
| s10b1028 | 3562.587       | 9487.505     | -484.794 | 405.496  | 12169.960 | -87.817 |
| .....    |                |              |          |          |           |         |
|          | Varianza total | 1724312.5534 |          |          |           |         |

|                                        |            |            |            |           |           |           |           |
|----------------------------------------|------------|------------|------------|-----------|-----------|-----------|-----------|
| Autovalores de la matriz de covarianza |            |            |            |           |           |           |           |
|                                        | Autovalor  | Diferencia | Proporción | Acumulada |           |           |           |
| 1                                      | 1294560.07 | 1155843.09 | 0.7508     | 0.7508    |           |           |           |
| 2                                      | 138716.98  | 76995.67   | 0.0804     | 0.8312    |           |           |           |
| 3                                      | 61721.31   | 18168.64   | 0.0358     | 0.8670    |           |           |           |
| 4                                      | 43552.67   | 5166.68    | 0.0253     | 0.8923    |           |           |           |
| 5                                      | 38385.99   | 10208.10   | 0.0223     | 0.9145    |           |           |           |
| 6                                      | 28177.89   | 3164.54    | 0.0163     | 0.9309    |           |           |           |
| 7                                      | 25013.35   | 11130.84   | 0.0145     | 0.9454    |           |           |           |
| 8                                      | 13882.51   | 2542.48    | 0.0081     | 0.9534    |           |           |           |
| 9                                      | 11340.03   | 1787.14    | 0.0066     | 0.9600    |           |           |           |
| 10                                     | 9552.89    | 2269.18    | 0.0055     | 0.9655    |           |           |           |
| 11                                     | 7283.70    | 34.85      | 0.0042     | 0.9698    |           |           |           |
| 12                                     | 7248.86    | 755.02     | 0.0042     | 0.9740    |           |           |           |
| 13                                     | 6493.84    | 231.53     | 0.0038     | 0.9777    |           |           |           |
| 14                                     | 6262.31    | 704.36     | 0.0036     | 0.9814    |           |           |           |
| 15                                     | 5557.95    | 855.57     | 0.0032     | 0.9846    |           |           |           |
| 16                                     | 4702.39    | 733.47     | 0.0027     | 0.9873    |           |           |           |
| 17                                     | 3968.92    | 548.00     | 0.0023     | 0.9896    |           |           |           |
| 18                                     | 3420.92    | 396.54     | 0.0020     | 0.9916    |           |           |           |
| 19                                     | 3024.39    | 204.73     | 0.0018     | 0.9934    |           |           |           |
| 20                                     | 2819.65    | 332.56     | 0.0016     | 0.9950    |           |           |           |
| 21                                     | 2487.09    | 836.22     | 0.0014     | 0.9964    |           |           |           |
| 22                                     | 1650.87    | 283.20     | 0.0010     | 0.9974    |           |           |           |
| 23                                     | 1367.67    | 345.90     | 0.0008     | 0.9982    |           |           |           |
| 24                                     | 1021.77    | 183.90     | 0.0006     | 0.9988    |           |           |           |
| 25                                     | 837.87     | 98.53      | 0.0005     | 0.9993    |           |           |           |
| 26                                     | 739.33     | 218.02     | 0.0004     | 0.9997    |           |           |           |
| 27                                     | 521.32     |            | 0.0003     | 1.0000    |           |           |           |
| Autovectores                           |            |            |            |           |           |           |           |
|                                        | Prin1      | Prin2      | Prin3      | Prin4     | Prin5     | Prin6     | Prin7     |
| s10b1001                               | 0.002910   | 0.027990   | 0.050795   | 0.004832  | 0.029579  | 0.051207  | 0.039379  |
| s10b1002                               | 0.007798   | 0.078646   | 0.061615   | 0.025166  | 0.459031  | 0.617356  | 0.446917  |
| s10b1003                               | -0.000382  | -0.000317  | 0.003243   | 0.013880  | 0.010456  | -0.028152 | 0.025728  |
| s10b1004                               | 0.000497   | 0.037955   | 0.025375   | 0.010209  | 0.065580  | 0.098914  | 0.079466  |
| s10b1005                               | 0.009685   | 0.044975   | 0.055208   | -0.012566 | 0.066213  | 0.048249  | 0.023686  |
| s10b1006                               | -0.000043  | 0.004243   | 0.007529   | 0.006208  | 0.016625  | 0.012468  | 0.014556  |
| s10b1007                               | 0.001762   | 0.031019   | 0.047305   | 0.010092  | 0.039413  | 0.045234  | 0.037477  |
| s10b1008                               | 0.001612   | 0.008803   | 0.021643   | 0.001730  | 0.008698  | 0.037910  | 0.022906  |
| s10b1009                               | 0.000666   | 0.019803   | 0.032182   | -0.006019 | 0.026944  | 0.014406  | 0.009653  |
| s10b1010                               | 0.015403   | 0.188839   | 0.609437   | -0.753853 | -0.114607 | 0.006825  | -0.035220 |
| s10b1011                               | 0.039865   | 0.949709   | -0.280549  | 0.030502  | -0.074655 | -0.062991 | -0.032228 |
| s10b1012                               | 0.000192   | 0.030950   | 0.001841   | 0.001204  | 0.023734  | 0.031773  | 0.024340  |
| s10b1013                               | 0.002738   | 0.098686   | 0.065592   | 0.046018  | 0.080221  | 0.170460  | 0.132471  |
| s10b1014                               | 0.006925   | 0.017686   | 0.067094   | 0.028976  | 0.119489  | -0.080308 | -0.070329 |
| s10b1015                               | 0.000511   | 0.081235   | 0.229303   | 0.140707  | 0.694954  | -0.286677 | -0.421150 |
| s10b1016                               | 0.000119   | 0.004905   | 0.028157   | 0.004911  | 0.041242  | -0.019173 | -0.023608 |
| s10b1017                               | 0.002228   | 0.031734   | 0.097721   | 0.048806  | 0.240180  | -0.149687 | -0.151553 |
| s10b1018                               | 0.001822   | 0.039826   | 0.072553   | 0.020825  | 0.124868  | 0.016083  | 0.013332  |
| s10b1019                               | 0.000536   | 0.014618   | 0.020513   | 0.002306  | 0.023231  | 0.006040  | 0.002263  |
| s10b1020                               | 0.000873   | 0.023592   | 0.070831   | 0.021547  | 0.056738  | -0.043031 | 0.009729  |
| s10b1021                               | 0.000087   | 0.018068   | 0.038904   | 0.009903  | 0.047771  | 0.006705  | 0.000730  |
| s10b1022                               | -0.000159  | 0.008813   | 0.073895   | 0.012781  | 0.107105  | -0.658325 | 0.733034  |
| s10b1024                               | 0.004346   | 0.013256   | 0.018478   | 0.000425  | 0.022699  | -0.015349 | 0.028188  |





|          |          |          |          |          |          |          |          |
|----------|----------|----------|----------|----------|----------|----------|----------|
| s10b1025 | 0.007268 | 0.136839 | 0.648313 | 0.630410 | -.365488 | 0.053125 | 0.005364 |
| s10b1026 | 0.003333 | 0.078240 | 0.152982 | 0.074958 | 0.160579 | 0.117273 | 0.152356 |
| s10b1027 | 0.003028 | 0.046290 | 0.050035 | -.009791 | 0.040875 | 0.040040 | 0.032021 |
| s10b1028 | 0.998923 | -.044113 | -.006156 | 0.004928 | -.000020 | -.003604 | -.002151 |

.....

## Comentario

Los resultados de la Tabla 38, presenta estadística descriptiva (media y desviación estándar); la matriz de covarianza; autovalores donde el primer componente la variación explicada esta 75.08 y el segundo 8.04 %; los autovectores de los componentes principales por variable, donde los valores próximos a cero poseen escasa variación.

En la Tabla 39, se incluye el procedimiento SORT y GPLOT, la primera para ordenar por departamento; la segunda para graficar el primer y segundo componente principal por cluster y en otra gráfica el componente tres y cuatro. Obsérvese la dispersión de los cluster en ambas (Tabla 40).

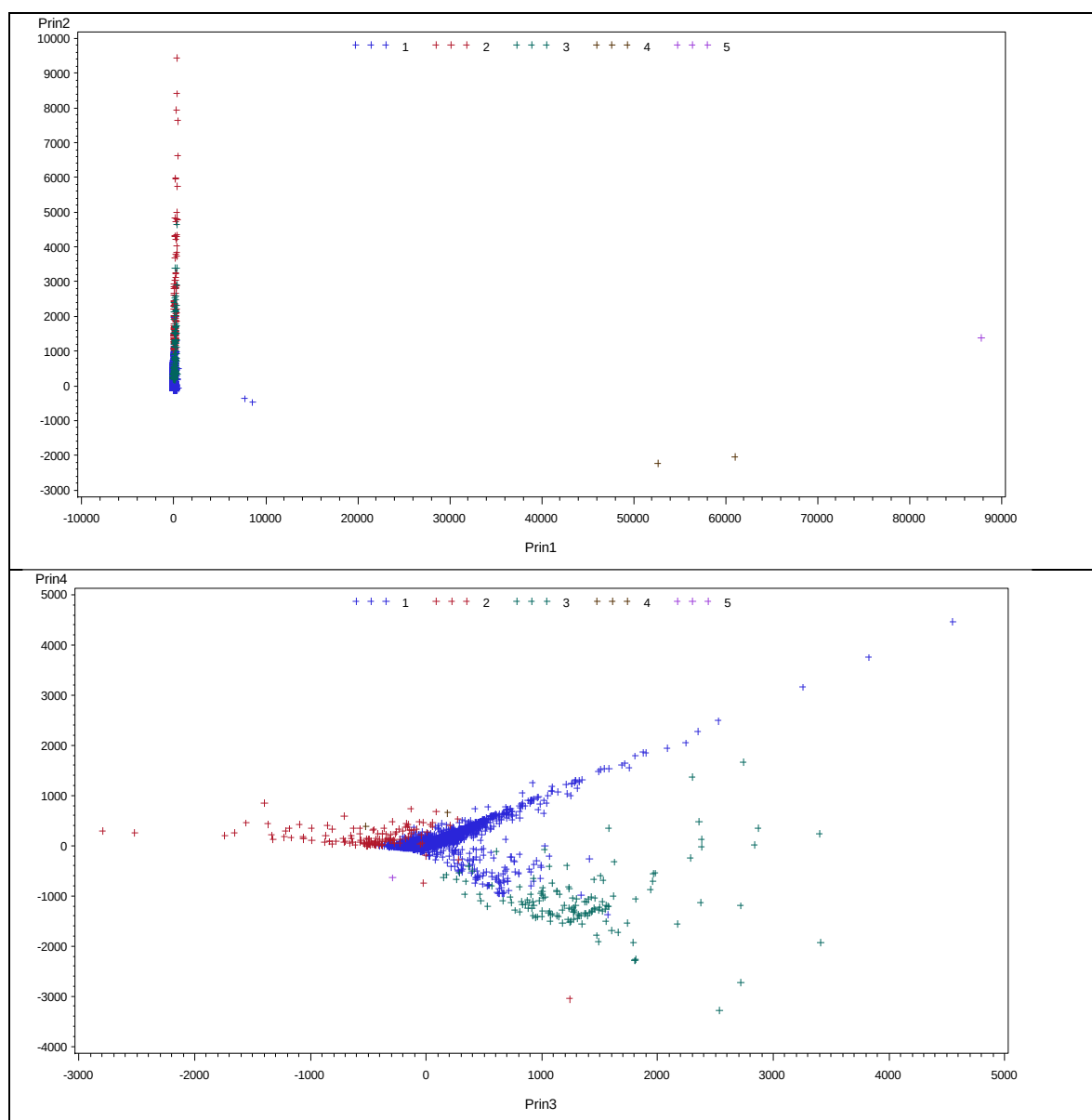
Tabla 39. Procedimiento SORT y GPLOT con sentencias para graficar los clusters en función a los componentes principales 1-2 y 3-4.

```
ods graphics off;
goptions reset=all border;
legend1 label= none
      position=(top center inside)
      mode=share;

proc sort data=new1; by depto; run;
proc gplot data=new1 ;
plot prin2*prin1=cluster / legend=legend1 ;
run;

proc gplot data=new1 ;
plot prin4*prin3=cluster / legend=legend1 ;
run;
quit;
```

Tabla 40. Figuras de los clusters en función a los componentes principales 1-2 y 3-4.



### Cluster Método Average (Variables cualitativas)

En el enlace promedio, la distancia entre dos grupos es la distancia promedio entre pares de observaciones, una en cada grupo. El enlace promedio tiende a unir grupos con pequeñas variaciones, y está ligeramente sesgado hacia la producción de grupos con la misma variación.

En base a la Evaluación de la campaña verano 2017-2018 y pronóstico de la campaña invierno 2018 del cultivo de papa (OAP-MDRyT, 2018), se obtuvo promedios de las variables para generar el archivo “papas epally programa.sas” (Tabla 41). En la Tabla 42, se expone el procedimiento FACTOR con el fin de determinar la variación explicada de los componentes principales y con las variables que poseen



mayor variación utiliza el procedimiento CLUSTER método AVERAGE, también genera un nuevo data sas para graficar el dendrograma. El procedimiento TREE permitirá obtener el dendrograma y otro data sas. Con el procedimiento SORT se ordena los data sas generadas por TREE con data sas inicial por el código de identificación denominada “folio”. Juntar los data sas con el procedimiento DATA con la sentencia MERGE y BY. Con el procedimiento MEANS se puede obtener información de variables cuantitativas por cluster. Los resultados se observan en la Tabla 43.

Tabla 41. Procedimiento y sentencia de DATA para datos de producción de papa.

```
data edgar;
input p01dep p02pro p03mun folio sup1718 rend1718 p16ccalf semt p18msem p19mcos
p20fcli1 p20plg1 p20enf1 p21proaf p23agrqm p24guan p24barr p24zanj p24aver p24comp
p24rot p24labr p24terr p24otr p25cods p26dest p27apoy p30rieg;
title1 ' ';
cards;
2 201 20201 114 0.01 10.00 2 3.45 11 5 5 98 98 20 1 1 1 2 1 2 2 2 2 2 2 5 2 1
.....
7 713 71301 191 1.00 19.20 2 3.12 5 9 2 98 98 15 1 2 2 2 1 2 2 2 2 2 2 2 2 1
;

options nodate ps=400;
```

Tabla 42. Procedimiento y sentencia de FACTOR, CLUSTER, TREE, SORT, DATA y MEANS para datos de producción de papa.

```
* reducción de variables ;
proc factor data=edgar nfactors=1 out=facsp04;
var p16ccalf p18msem p19mcos p20fcli1 p23agrqm p24guan p24barr
p24zanj p24aver p24comp p24rot p24labr p24terr p24otr p25cods p26dest
p27apoy p30rieg;
run;

* cluster metodo average;

proc cluster method=ave data=edgar outtree=tree03 ccc pseudo ;
var p18msem p23agrqm p24guan p24barr p24aver p24comp p24labr
p24terr p24otr p25cods p26dest p27apoy p30rieg;
id folio;
run;

* figura de dendrograma;

proc tree data=tree03 n=5 HORDISPLAY=RIGHT out=arb;
id folio;
run;
quit;

* promedios de rendimiento de papa por grupo;
proc sort data= arb; by folio; run;
```



```
proc sort data=edgar; by folio; run;
data pedgar; merge arb edgar; by folio; run;

proc means data=pedgar maxdec=2;
class cluster;
var rend1718;
run;
```

Tabla 43. Salidas del procedimiento y sentencia de FACTOR, CLUSTER, TREE, SORT, DATA y MEANS, para datos de producción de papa.

|                                                                  |            |            |            |            |            |         |
|------------------------------------------------------------------|------------|------------|------------|------------|------------|---------|
| Procedimiento FACTOR                                             |            |            |            |            |            |         |
| Método factorial inicial Componentes principales                 |            |            |            |            |            |         |
| Estimaciones de comunalidad anterior: ONE                        |            |            |            |            |            |         |
| Autovalores de la matriz de correlación: Total = 18 Promedio = 1 |            |            |            |            |            |         |
|                                                                  | Autovalor  | Diferencia | Proporción | Acumulada  |            |         |
| 1                                                                | 2.47002214 | 0.79125445 | 0.1372     | 0.1372     |            |         |
| 2                                                                | 1.67876769 | 0.23298226 | 0.0933     | 0.2305     |            |         |
| 3                                                                | 1.44578543 | 0.20216149 | 0.0803     | 0.3108     |            |         |
| 4                                                                | 1.24362394 | 0.06007243 | 0.0691     | 0.3799     |            |         |
| 5                                                                | 1.18355150 | 0.10765681 | 0.0658     | 0.4457     |            |         |
| 6                                                                | 1.07589470 | 0.02650990 | 0.0598     | 0.5054     |            |         |
| 7                                                                | 1.04938480 | 0.04351312 | 0.0583     | 0.5637     |            |         |
| 8                                                                | 1.00587168 | 0.08123252 | 0.0559     | 0.6196     |            |         |
| 9                                                                | 0.92463916 | 0.03217632 | 0.0514     | 0.6710     |            |         |
| 10                                                               | 0.89246284 | 0.02029120 | 0.0496     | 0.7206     |            |         |
| 11                                                               | 0.87217164 | 0.09331349 | 0.0485     | 0.7690     |            |         |
| 12                                                               | 0.77885816 | 0.05303349 | 0.0433     | 0.8123     |            |         |
| 13                                                               | 0.72582466 | 0.01968344 | 0.0403     | 0.8526     |            |         |
| 14                                                               | 0.70614122 | 0.07840607 | 0.0392     | 0.8918     |            |         |
| 15                                                               | 0.62773515 | 0.11710631 | 0.0349     | 0.9267     |            |         |
| 16                                                               | 0.51062885 | 0.07307843 | 0.0284     | 0.9551     |            |         |
| 17                                                               | 0.43755041 | 0.06646440 | 0.0243     | 0.9794     |            |         |
| 18                                                               | 0.37108601 | 0.0206     | 1.0000     |            |            |         |
| El factor 1 estará retenido por el criterio NFACTOR.             |            |            |            |            |            |         |
| Modelo factorial                                                 |            |            |            |            |            |         |
| Factor1                                                          |            |            |            |            |            |         |
|                                                                  | p16ccalf   | -0.06192   |            |            |            |         |
|                                                                  | p18msem    | -0.31903   |            |            |            |         |
|                                                                  | p19mcos    | 0.14509    |            |            |            |         |
|                                                                  | p20fcli1   | 0.13922    |            |            |            |         |
|                                                                  | p23agrqm   | -0.41294   |            |            |            |         |
|                                                                  | p24guan    | 0.49941    |            |            |            |         |
|                                                                  | p24barr    | 0.30413    |            |            |            |         |
|                                                                  | p24zanj    | -0.01265   |            |            |            |         |
|                                                                  | p24aver    | 0.21555    |            |            |            |         |
|                                                                  | p24comp    | 0.73104    |            |            |            |         |
|                                                                  | p24rot     | 0.13850    |            |            |            |         |
|                                                                  | p24labr    | 0.63064    |            |            |            |         |
|                                                                  | p24terr    | 0.31428    |            |            |            |         |
|                                                                  | p24otr     | -0.27033   |            |            |            |         |
|                                                                  | p25cods    | 0.68845    |            |            |            |         |
|                                                                  | p26dest    | -0.23941   |            |            |            |         |
|                                                                  | p27apoy    | -0.26165   |            |            |            |         |
|                                                                  | p30rieg    | -0.20499   |            |            |            |         |
| Varianza explicada por cada factor                               |            |            |            |            |            |         |
| Factor1                                                          |            |            |            |            |            |         |
| 2.4700221                                                        |            |            |            |            |            |         |
| Estimaciones de comunalidad final: total = 2.470022              |            |            |            |            |            |         |
|                                                                  | p16ccalf   | p18msem    | p19mcos    | p20fcli1   | p23agrqm   | p24guan |
| 0.00383403                                                       | 0.10178066 | 0.02105032 | 0.01938176 | 0.17052090 | 0.24941310 |         |
|                                                                  | p24barr    | p24zanj    | p24aver    | p24comp    | p24rot     | p24labr |
| 0.09249605                                                       | 0.00016003 | 0.04645998 | 0.53441780 | 0.01918327 | 0.39770510 |         |
|                                                                  | p24terr    | p24otr     | p25cods    | p26dest    | p27apoy    | p30rieg |
| 0.09877441                                                       | 0.07307918 | 0.47396534 | 0.05731625 | 0.06846205 | 0.04202192 |         |

Procedimiento FACTOR  
Método factorial inicial Componentes principales

### Coeficientes de puntuación estimados por la regresión

Correlaciones múltiples cuadradas de las variables con cada factor

Factor1

1.00000000

### Coeficientes de puntuación estandarizados

Factor1

|          |          |
|----------|----------|
| p16ccalf | -0.02507 |
| p18msem  | -0.12916 |
| p19mcos  | 0.05874  |
| p20fcfl1 | 0.05636  |
| p23agrgm | -0.16718 |
| p24guan  | 0.20219  |
| p24barr  | 0.12313  |
| p24zanj  | -0.00512 |
| p24aver  | 0.08726  |
| p24comp  | 0.29596  |
| p24rot   | 0.05607  |
| p24labr  | 0.25532  |
| p24terr  | 0.12724  |
| p24otr   | -0.10945 |
| p25cods  | 0.27872  |
| p26dest  | -0.09693 |
| p27apoy  | -0.10593 |
| p30rieg  | -0.08299 |

### Procedimiento CLUSTER

#### Análisis del conglomerado de enlace medio

Autovalores de la matriz de covarianza

| Autovalor | Diferencia | Proporción | Acumulada |
|-----------|------------|------------|-----------|
|-----------|------------|------------|-----------|

|    |            |            |        |        |
|----|------------|------------|--------|--------|
| 1  | 4.46132626 | 2.69859639 | 0.5915 | 0.5915 |
| 2  | 1.27627987 | 1.46681486 | 0.2337 | 0.8251 |
| 3  | 0.94641502 | 0.06627222 | 0.0391 | 0.8642 |
| 4  | 0.22798779 | 0.04669017 | 0.0302 | 0.8944 |
| 5  | 0.18129762 | 0.03015306 | 0.0240 | 0.9185 |
| 6  | 0.15114456 | 0.01473661 | 0.0200 | 0.9385 |
| 7  | 0.13640794 | 0.01825265 | 0.0181 | 0.9566 |
| 8  | 0.11455529 | 0.01614756 | 0.0152 | 0.9718 |
| 9  | 0.09840774 | 0.05000365 | 0.0130 | 0.9848 |
| 10 | 0.04840409 | 0.00872876 | 0.0064 | 0.9912 |
| 11 | 0.03967533 | 0.01557307 | 0.0053 | 0.9965 |
| 12 | 0.02410226 | 0.02173975 | 0.0032 | 0.9997 |
| 13 | 0.00236251 |            | 0.0003 | 1.0000 |

Desviación estándar muestral total de la raíz cuadrada de la media 0.76173

Distancia de la raíz cuadrada media entre observaciones 3.884074

## Historia de conglomerado

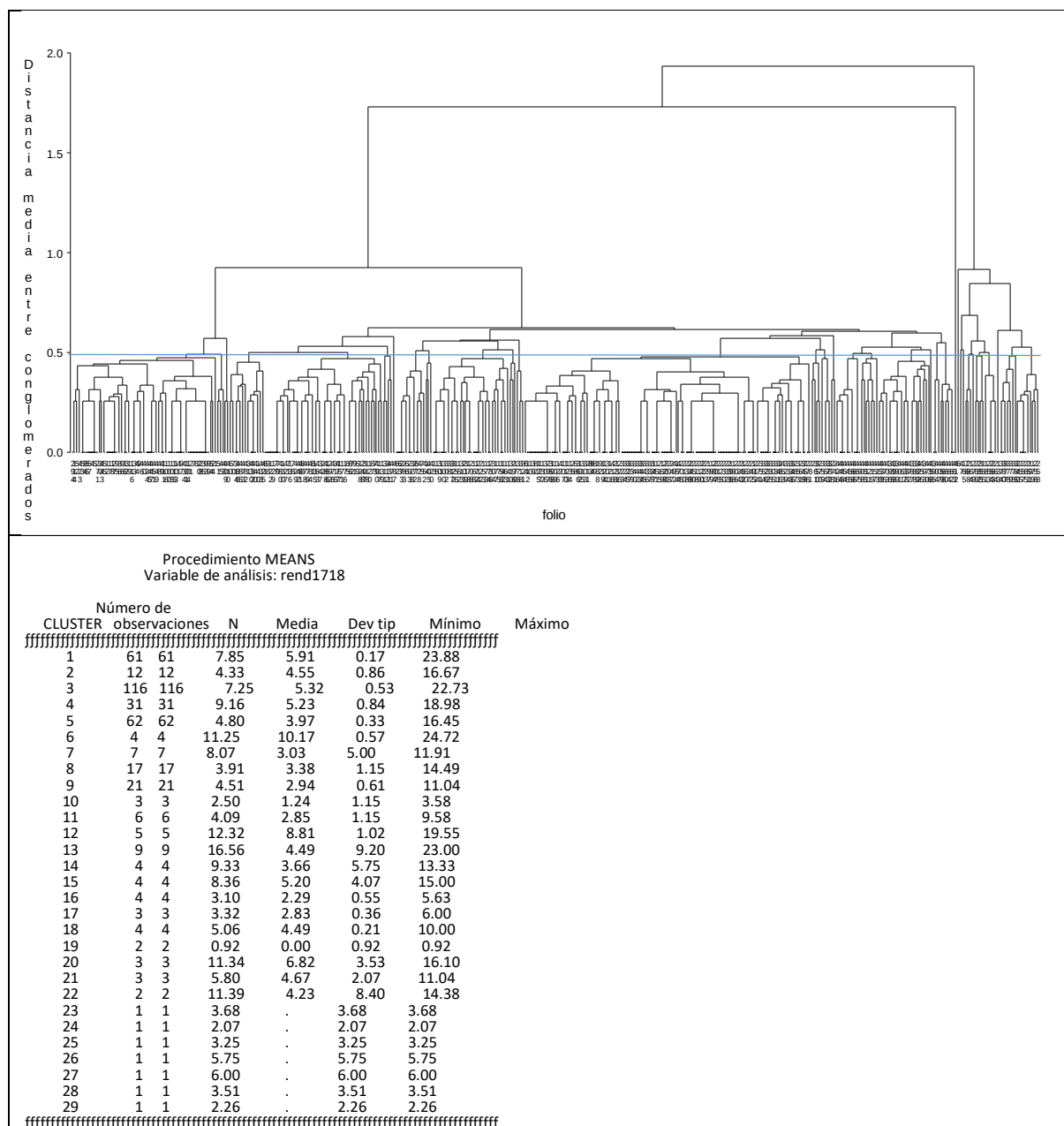
Norm T  
RMS i

|     |        |     |   |        |      |   |   |   |   |   |   |   |
|-----|--------|-----|---|--------|------|---|---|---|---|---|---|---|
| 389 | 56     | 57  | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 388 | 62     | 63  | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 387 | 100    | 105 | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 386 | 108    | 109 | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 385 | 107    | 110 | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 384 | 49     | 111 | 2 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |
| 383 | Cl 387 | 115 | 3 | 0.0000 | 1.00 | . | . | . | . | . | 0 | T |

• • • • •

.....  
15

|    |      |       |     |        |      |      |      |     |        |        |
|----|------|-------|-----|--------|------|------|------|-----|--------|--------|
| 16 | CL18 | CL62  | 52  | 0.0022 | .803 | .798 | 0.95 | 102 | 3.7    | 0.5602 |
| 15 | CL32 | CL174 | 63  | 0.0022 | .801 | .793 | 1.37 | 108 | 4.7    | 0.5695 |
| 14 | CL36 | CL25  | 124 | 0.0055 | .795 | .788 | 1.25 | 112 | 11.3   | 0.5728 |
| 13 | CL20 | 487   | 67  | 0.0012 | .794 | .783 | 1.73 | 121 | 2.2    | 0.58   |
| 12 | CL14 | CL81  | 126 | 0.0021 | .792 | .777 | 2.32 | 131 | 4.0    | 0.5822 |
| 11 | CL21 | CL19  | 48  | 0.0058 | .786 | .770 | 2.44 | 139 | 10.1   | 0.5965 |
| 10 | CL12 | CL11  | 174 | 0.0235 | .763 | .762 | 0.11 | 136 | 40.7   | 0.6047 |
| 9  | 133  | CL22  | 17  | 0.0013 | .761 | .753 | 0.99 | 152 | 2.6    | 0.6106 |
| 8  | CL16 | CL10  | 226 | 0.0247 | .737 | .741 | -.51 | 153 | 35.7   | 0.6145 |
| 7  | CL13 | CL8   | 293 | 0.0349 | .702 | .727 | -2.6 | 150 | 47.2   | 0.6252 |
| 6  | CL23 | CL17  | 15  | 0.0026 | .699 | .708 | -.93 | 178 | 4.1    | 0.6849 |
| 5  | CL6  | CL9   | 32  | 0.0193 | .680 | .683 | -.27 | 204 | 29.0   | 0.8473 |
| 4  | 6    | CL5   | 33  | 0.0030 | .677 | .644 | 3.21 | 269 | 2.4    | 0.9171 |
| 3  | CL15 | CL7   | 356 | 0.1585 | .518 | .581 | -5.3 | 208 | 200    | 0.9242 |
| 2  | CL3  | 412   | 357 | 0.0142 | .504 | .445 | 0.12 | 395 | 11.4   | 1.7321 |
| 1  | CL2  | CL4   | 390 | 0.5042 | .000 | .000 | 0.00 | 395 | 19.347 |        |



## Comentario

El procedimiento FACTOR (Tabla 43) presenta variación explicada en el primer componente principal 13.72 % el segundo 9.33 % y disminuye en los siguientes componentes. Así mismo las variables p16ccalf, p19mcos, p20fclii, p24zanj y p24rot por valores absolutos menores a 0.06 los coeficientes de puntuación estandarizados se excluyeron.

El procedimiento CLUSTER método AVERAGE presenta una variación explicada en el primer componente principal un 59.15 % y el segundo un 23.37% sumados ambos 82.51 % explican sobre la variación total. El número de grupos o clusters a una distancia media de entre conglomerados de 0.5 son 29, y se visualiza en el dendrograma con un alineamiento divisor. Con una estadística descriptiva de



variables cuantitativa como el rendimiento se caracterizan los clusters. Clasificando con este procedimiento a los encuestados e identificando por su afinidad se generaría estrategias para el desarrollo de los encuestados.

#### 4.4. Análisis de regresión logística

El análisis de regresión logística es un método explicativo de un fenómeno dicotómico por variables independientes categóricos o cuantitativos

Otras características del procedimiento LOGÍSTICO le permiten hacer lo siguiente:

- Controlar el orden de las categorías de respuesta
- Calcule una medida generalizada de  $R^2$  para el modelo ajustado
- Reclasificar las observaciones de respuesta binaria de acuerdo con sus probabilidades de respuesta predichas
- Probar hipótesis lineales sobre los parámetros de regresión.
- Crear un conjunto de datos para producir una curva característica de funcionamiento del receptor para cada modelo ajustado
- Especificar contrastes para comparar varias curvas características de funcionamiento del receptor.
- Crear un conjunto de datos que contenga las probabilidades de respuesta estimadas, los residuos y los diagnósticos de influencia

En base a la Evaluación de la campaña verano 2017-2018 y pronóstico de la campaña invierno 2018 del cultivo de papa (OAP-MDRyT, 2018), se obtuvo promedios de las variables para generar el archivo “papas epally programa.sas” (Tabla 41). En dicha data sas, el rendimiento fue categorizado, si fue mayor al indicador (5.98 t/ha) con 1 y si fue menor al indicador con 0, denominándose la nueva variable “rendcla”. En la Tabla 44 se presenta el procedimiento LOGISTIC con sentencias sobre el rendimiento para responder ¿cómo influyen las covariables?. Los resultados se exponen en la Tabla 45.

Tabla 44. Procedimiento LOGISTIC con sentencias para determinar el rendimiento (categorizado) como es afectada por las covariables, con y sin selección por el método Stepwise.

```
proc logistic data=edgar;
model rendcla = sup1718 semt p18msem p19mcos p20plg1 p20enf1 p21proaf
/EXPB selection=stepwise;
output out=esti predicted=ocu;
run;

proc print data=esti;
run;

proc logistic data=edgar;
```



```
model rendcla = sup1718 semt p18msem p19mcos p20plg1 p20enf1 p21proaf
/EXPB ;
output out=esti predicted=ocu;
run;
```

Tabla 45. Resultados del procedimiento LOGISTIC con sentencias para determinar el rendimiento (categorizado) como es afectada por las covariables, con y sin selección por el método Stepwise.

|                                                                                                    |                              |                         |            |
|----------------------------------------------------------------------------------------------------|------------------------------|-------------------------|------------|
| Procedimiento LOGISTIC                                                                             |                              |                         |            |
| Información del modelo                                                                             |                              |                         |            |
| Conjunto de datos                                                                                  | WORK.EDGAR                   |                         |            |
| Variable de respuesta                                                                              | rendcla                      |                         |            |
| Número de niveles de respuesta                                                                     | 2                            |                         |            |
| Modelo                                                                                             | logit binario                |                         |            |
| Técnica de optimización                                                                            | Puntuación de Fisher         |                         |            |
|                                                                                                    |                              |                         |            |
| Número de observaciones leí                                                                        | 391                          |                         |            |
| Número de observaciones usa                                                                        | 384                          |                         |            |
|                                                                                                    |                              |                         |            |
| Perfil de respuesta                                                                                |                              |                         |            |
| Valor ordenado                                                                                     | Frecuencia rendcla           | total                   |            |
| 1                                                                                                  | 0                            | 227                     |            |
| 2                                                                                                  | 1                            | 157                     |            |
|                                                                                                    |                              |                         |            |
| La probabilidad modelada es rendcla=0.                                                             |                              |                         |            |
|                                                                                                    |                              |                         |            |
| NOTA: 7 observations were deleted due to missing values for the response or explanatory variables. |                              |                         |            |
|                                                                                                    |                              |                         |            |
| Procedimiento de selección stepwise                                                                |                              |                         |            |
| Paso 0. Término independiente introducido:                                                         |                              |                         |            |
|                                                                                                    |                              |                         |            |
| Estado de convergencia del modelo                                                                  |                              |                         |            |
| Criterio de convergencia (GCONV=1E-8) satisfecho.                                                  |                              |                         |            |
| -2 LOG L = 519.505                                                                                 |                              |                         |            |
|                                                                                                    |                              |                         |            |
| Test residual de chi-cuadrado                                                                      |                              |                         |            |
| Chi-cuadrado                                                                                       | DF                           | Pr > ChiSq              |            |
| 70.8166                                                                                            | 7                            | <.0001                  |            |
|                                                                                                    |                              |                         |            |
| Paso 1. Efecto semt introducido:                                                                   |                              |                         |            |
|                                                                                                    |                              |                         |            |
| Estado de convergencia del modelo                                                                  |                              |                         |            |
| Criterio de convergencia (GCONV=1E-8) satisfecho.                                                  |                              |                         |            |
|                                                                                                    |                              |                         |            |
| Estadístico de ajuste del modelo                                                                   |                              |                         |            |
|                                                                                                    | Términos independientes      |                         |            |
|                                                                                                    | Sólo términos independientes | y Variables adicionales |            |
| Criterio                                                                                           |                              |                         |            |
| AIC                                                                                                | 521.505                      | 454.465                 |            |
| SC                                                                                                 | 525.456                      | 462.367                 |            |
| -2 LOG L                                                                                           | 519.505                      | 450.465                 |            |
|                                                                                                    |                              |                         |            |
| Probar hipótesis nula global: BETA=0                                                               |                              |                         |            |
| Test                                                                                               | Chi-cuadrado                 | DF                      | Pr > ChiSq |
| Ratio de verosim                                                                                   | 69.0396                      | 1                       | <.0001     |
| Puntuación                                                                                         | 65.4087                      | 1                       | <.0001     |
| Wald                                                                                               | 54.9550                      | 1                       | <.0001     |
|                                                                                                    |                              |                         |            |
| Test residual de chi-cuadrado                                                                      |                              |                         |            |





**Fortalecimiento de la capacidad institucional en los sectores de desarrollo integral con coca, tráfico ilícito de drogas y seguridad alimentaria para una eficiente gestión del apoyo presupuestario sectorial – Contrato n° DCI/LA/2017/392-699**



|              |  |    |            |  |
|--------------|--|----|------------|--|
| Chi-cuadrado |  | DF | Pr > ChiSq |  |
| 6.2165       |  | 6  | 0.3994     |  |

NOTA: No se han eliminado efectos para el modelo en el Paso 1.

NOTA: No (additional) effects met the 0.05 significance level for entry into the model.

Resumen de selección de paso a paso

| Paso | Efecto      | Número    | Chi-cuadrado | Chi-cuadrado     |            |
|------|-------------|-----------|--------------|------------------|------------|
|      | Introducido | Eliminado | DF           | en de puntuación | de Wald    |
|      |             |           |              |                  | Pr > ChiSq |
| 1    | semt        | 1         | 1            | 65.4087          | <.0001     |

Análisis del estimador de máxima verosimilitud

| Parámetro | DF | Error     | Chi-cuadrado |         |            |          |
|-----------|----|-----------|--------------|---------|------------|----------|
|           |    | Estimador | estándar     | de Wald | Pr > ChiSq | Exp(Est) |
| Intercept | 1  | 1.8357    | 0.2289       | 64.3085 | <.0001     | 6.269    |
| semt      | 1  | -1.3578   | 0.1832       | 54.9550 | <.0001     | 0.257    |

Estimadores de cocientes de disparidad;

| Efecto | Estimador del punto | 95% Wald | Límites de confianza |
|--------|---------------------|----------|----------------------|
| semt   | 0.257               | 0.180    | 0.368                |

Asociación de probabilidades predichas y respuestas observadas

|                            |       |             |       |
|----------------------------|-------|-------------|-------|
| Concordancia de porcentaje | 72.2  | D de Somers | 0.469 |
| Discordancia de porcentaje | 25.3  | Gamma       | 0.481 |
| Porcentaje ligado          | 2.5   | Tau-a       | 0.227 |
| Pares                      | 35639 | c           | 0.735 |

Procedimiento LOGISTIC

Información del modelo

|                                |                      |
|--------------------------------|----------------------|
| Conjunto de datos              | WORK.EDGAR           |
| Variable de respuesta          | rendcla              |
| Número de niveles de respuesta | 2                    |
| Modelo                         | logit binario        |
| Técnica de optimización        | Puntuación de Fisher |

Número de observaciones leí

Número de observaciones usa

391

384

Perfil de respuesta

| Valor ordenado | Frecuencia |       |
|----------------|------------|-------|
|                | rendcla    | total |
| 1              | 0          | 227   |
| 2              | 1          | 157   |

La probabilidad modelada es rendcla=0.

NOTA: 7 observations were deleted due to missing values for the response or explanatory variables.

Estado de convergencia del modelo

Criterio de convergencia (GCONV=1E-8) satisfecho.

Estadístico de ajuste del modelo

| Criterio | Sólo términos independientes | Términos independientes y Variables adicionales |
|----------|------------------------------|-------------------------------------------------|
| AIC      | 521.505                      | 459.974                                         |
| SC       | 525.456                      | 491.579                                         |
| -2 LOG L | 519.505                      | 443.974                                         |

Probar hipótesis nula global: BETA=0

| Test             | Chi-cuadrado | DF | Pr > ChiSq |
|------------------|--------------|----|------------|
| Ratio de verosim | 75.5309      | 7  | <.0001     |
| Puntuación       | 70.8166      | 7  | <.0001     |
| Wald             | 58.4777      | 7  | <.0001     |



| Análisis del estimador de máxima verosimilitud |    |                 |                       |         |            |          |
|------------------------------------------------|----|-----------------|-----------------------|---------|------------|----------|
| Parámetro                                      | DF | Error Estimator | Chi-cuadrado estándar | de Wald | Pr > ChiSq | Exp(Est) |
| Intercept                                      | 1  | 0.7572          | 0.7120                | 1.1311  | 0.2875     | 2.132    |
| sup1718                                        | 1  | 0.0759          | 0.0891                | 0.7247  | 0.3946     | 1.079    |
| semt                                           | 1  | -1.3364         | 0.1858                | 51.7065 | <.0001     | 0.263    |
| p18mse                                         | 1  | 0.0778          | 0.0576                | 1.8231  | 0.1769     | 1.081    |
| p19mcos                                        | 1  | -0.0269         | 0.0351                | 0.5871  | 0.4435     | 0.973    |
| p20plg1                                        | 1  | 0.00264         | 0.00316               | 0.6978  | 0.4035     | 1.003    |
| p20enf1                                        | 1  | 0.00222         | 0.00357               | 0.3870  | 0.5339     | 1.002    |
| p21proaf                                       | 1  | 0.000959        | 0.00125               | 0.5861  | 0.4439     | 1.001    |

| Estimadores de cocientes de disparidad; |                     |          |                      |
|-----------------------------------------|---------------------|----------|----------------------|
| Efecto                                  | Estimador del punto | 95% Wald | Límites de confianza |
| sup1718                                 | 1.079               | 0.906    | 1.285                |
| semt                                    | 0.263               | 0.183    | 0.378                |
| p18mse                                  | 1.081               | 0.965    | 1.210                |
| p19mcos                                 | 0.973               | 0.909    | 1.043                |
| p20plg1                                 | 1.003               | 0.996    | 1.009                |
| p20enf1                                 | 1.002               | 0.995    | 1.009                |
| p21proaf                                | 1.001               | 0.999    | 1.003                |

| Asociación de probabilidades predichas y respuestas observadas |       |             |       |
|----------------------------------------------------------------|-------|-------------|-------|
| Concordancia de porcentaje                                     | 73.8  | D de Somers | 0.477 |
| Discordancia de porcentaje                                     | 26.0  | Gamma       | 0.478 |
| Porcentaje ligado                                              | 0.2   | Tau-a       | 0.231 |
| Pares                                                          | 35639 | c           | 0.739 |

## Comentario

En la Tabla 45, reporta 384 observaciones utilizadas, luego de la primera etapa la variable semilla empleada fue seleccionada, éste presenta un coeficiente de regresión de 0.257 t/ha a nivel muestral y es diferente de cero a nivel poblacional a P: 0.01.

En el segundo resultado la prueba de Wald para los coeficientes de regresión logística solo es significativa la variable semilla empleada a P: 0.01 deduciendo a nivel poblacional es diferente de cero. A nivel muestral se tiene un incremento de 0.263 t/ha. Con el procedimiento PRINT se visualiza la probabilidad de ocurrencia del encuestado en superar el indicador mayor a 5.98 t/ha.



## BIBLIOGRAFÍA

GUJARATI, D. 2002. Econometría. Trad. de tercera edición BASIC ECONOMETRICS. McGraw-Hill. Colombia pp 100-101.

LEON, C; QUIROZ, R. 1994. Análisis de sistemas agropecuarios; uso de métodos matemáticos. CIRNMA Centro Investigación de Recursos Naturales y Medio Ambiente. La Paz Bolivia. 236 p.

PEDRET R; SAGNIER L; CAMP F. 2003. Herramientas para segmentar mercados y posicionar productos. Análisis de información cuantitativa en investigación comercial. Ediciones DEUSTO. España.

PEREZ, C. 2001. El sistema estadístico SAS®. Prentice Hall. Impreso en España.

SAS INSTITUTE INC. 2004. Base SAS® 9.1 Procedures Guide. Cary, NC: SAS Institute Inc.

SAS INSTITUTE INC. 2004. SAS/STAT® 9.2 User's Guide Guide. Cary, NC: SAS Institute Inc.

SPIEGEL Y STEPHENS. 2009. Estadística. McGraw-Hill. Cuarta edición.

STEEL, TORRIE AND DICKEY. 1992. Bioestadística: Principios y procedimientos. McGraw-Hill.