

## Laboratorio 3 - Inferencia Estadística

### Laboratorio Semestral

#### Integrantes

- Gianfranco Astorga Saco
- JeanLucas Peñaloza
- Diego Godoy

## Fecha de Entrega

- 26 Noviembre del 2023

### 1. Introducción

El presente Laboratorio tiene como objetivo fundamental asesorar al gerente de desarrollo sustentable de SERVSAT mediante la aplicación de un modelo de regresión lineal. Este modelo buscará identificar y analizar las variables que inciden significativamente en la percepción de los trabajadores de la empresa SERVSAT, respecto a la Responsabilidad Social Empresarial (RSE) de la empresa, utilizando como base una muestra representativa de la población de trabajadores de la empresa SERVSAT.

Por otra parte, se pretende proporcionar una herramienta que permita a la empresa SERVSAT, identificar las variables que inciden significativamente en la percepción de los trabajadores respecto a la RSE de la empresa, con el fin de que la empresa pueda enfocar sus esfuerzos en mejorar las variables que inciden negativamente en la percepción de los trabajadores respecto a la RSE de la empresa.

#### 1.1 Descripción de la base de datos

La base de datos utilizada para el presente laboratorio corresponde a una muestra representativa de la población de trabajadores de la empresa SERVSAT, la cual fue obtenida mediante una encuesta aplicada a los trabajadores de la empresa SERVSAT. La encuesta fue aplicada a 100 trabajadores de la empresa SERVSAT, de los cuales 50 trabajadores pertenecen al área de operaciones y 50 trabajadores pertenecen al área de administración. La encuesta aplicada a los trabajadores de la empresa SERVSAT, consta de 20 preguntas, las cuales se encuentran relacionadas con la percepción de los trabajadores respecto a la RSE de la empresa SERVSAT.

### 2. Desarrollo

#### 2.1 Preparación de la base de datos

En primer lugar, se procede a cargar la base de datos en R, para ello se utiliza la función `foreign::read_spss`, la cual permite cargar archivos de extensión `.sav` en R.

```
# Cargar librerías
```

```
library(haven)
```

```
# Cargar base de datos
```

```
Percepcion_de_la_RSE <- read_sav("C:/Users/choco/OneDrive/Documentos/GitHub/Laboratorio-3-Inferencia-Es-
```

```
#View(Percepcion_de_la_RSE) #nolint
#head(Percepcion_de_la_RSE) #nolint
```

```
# Analisis descriptivo de las variables de la base de datos
```

```
summary(Percepcion_de_la_RSE)
```

```
##          ID              sexo          Edad          RAmbiental
## Min.   : 1.00    Min.   :1.000    Length:110    Min.   :1.000
## 1st Qu.: 28.25    1st Qu.:1.000    Class :character 1st Qu.:3.605
## Median : 55.50    Median :2.000    Mode  :character Median :4.593
## Mean   : 55.50    Mean   :1.673                      Mean   :4.547
## 3rd Qu.: 82.75    3rd Qu.:2.000                      3rd Qu.:5.427
## Max.   :110.00    Max.   :2.000                      Max.   :7.000
## REconómica    RSocial    SaludSeguridad    RSE
## Min.   :3.283    Min.   :1.000    Min.   :2.239    Min.   :1.952
## 1st Qu.:5.796    1st Qu.:3.397    1st Qu.:4.654    1st Qu.:4.310
## Median :6.172    Median :4.427    Median :5.470    Median :5.358
## Mean   :6.122    Mean   :4.339    Mean   :5.251    Mean   :5.118
## 3rd Qu.:6.828    3rd Qu.:5.186    3rd Qu.:6.000    3rd Qu.:6.000
## Max.   :7.000    Max.   :7.000    Max.   :7.000    Max.   :7.000
```

```
#nolint
```

### 3. Desarrollo

#### 3.1 Matriz de Correlaciones

En primer lugar, se procede a realizar una matriz de correlaciones entre las variables de la base de datos, para ello se utiliza la función cor, la cual permite obtener una matriz de correlaciones entre las variables de la base de datos.

```
# Matriz de correlaciones entre las variables de la base de datos
```

```
correlation_matrix <- cor(Percepcion_de_la_RSE[c("RSE", "RAmbiental", "REconómica", "RSocial", "SaludSegu
print(correlation_matrix)
```

```
##          RSE RAmbiental REconómica  RSocial SaludSeguridad
## RSE          1.0000000  0.7225707  0.6555541  0.6545916    0.6492963
## RAmbiental    0.7225707  1.0000000  0.4538625  0.6492129    0.6002047
## REconómica    0.6555541  0.4538625  1.0000000  0.3922374    0.4575324
## RSocial       0.6545916  0.6492129  0.3922374  1.0000000    0.6505262
## SaludSeguridad 0.6492963  0.6002047  0.4575324  0.6505262    1.0000000
```

```
#nolint
```

##### 3.1.1 Interpretación de la matriz de correlaciones

Se observa que la variable RSE presenta una correlación positiva y significativa con las variables RAmbiental, REconómica, RSocial y SaludSeguridad, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental, Responsabilidad Económica, Responsabilidad Social y Salud y Seguridad de la empresa.

### 3.2 Grafico de relación entre Responsabilidad Empresarial y Responsabilidad Ambiental

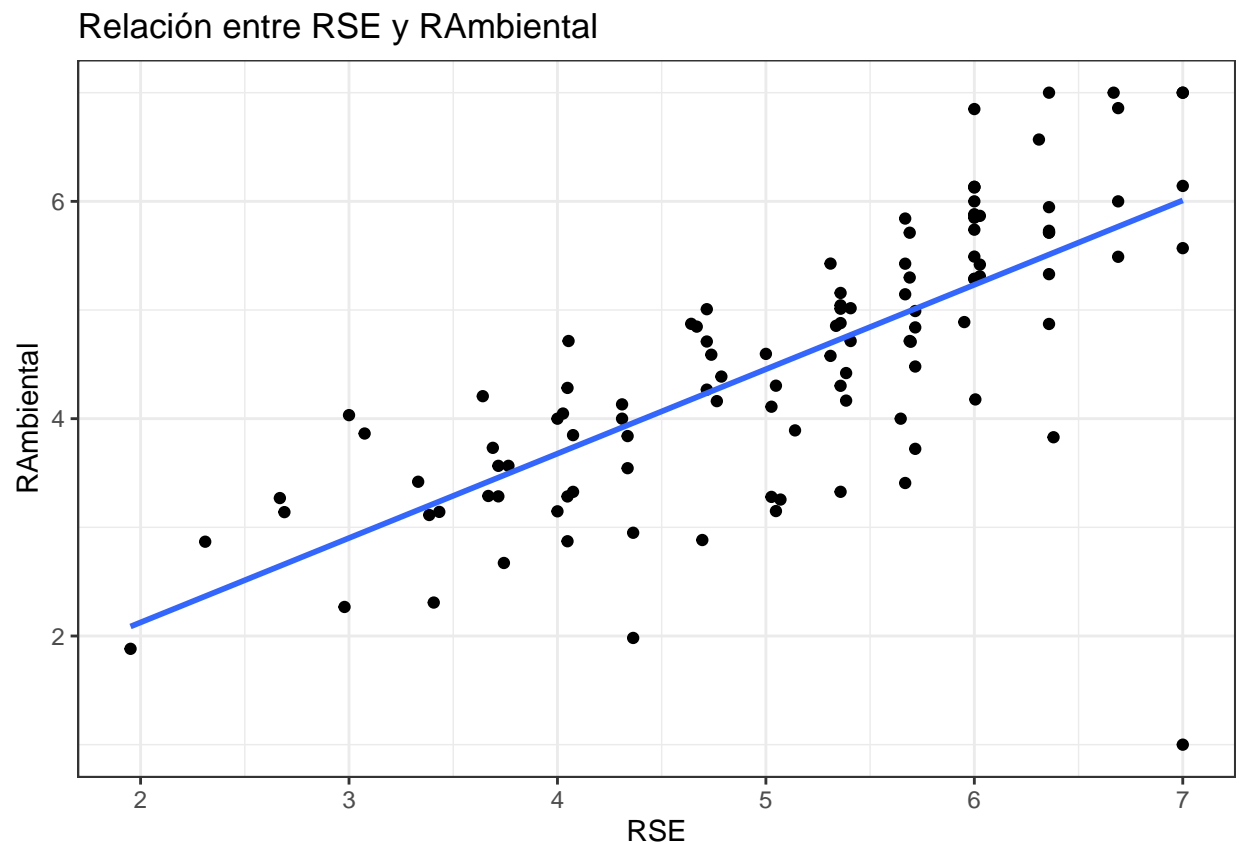
Se procede a graficar la relación entre la variable RSE y la variable RAmbiental, para ello se utiliza la función ggplot, la cual permite graficar la relación entre dos variables.

```
# Grafico de relacion entre RSE y RAmbiental
```

```
library(ggplot2)
```

```
ggplot(Percepcion_de_la_RSE, aes(x = RSE, y = RAmbiental)) + geom_point() + geom_smooth(method = "lm",
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
#nolint
```

#### 3.2.1 Interpretación del grafico de relación entre Responsabilidad Empresarial y Responsabilidad Ambiental

Se observa que existe una relación positiva y significativa entre la variable RSE y la variable RAmbiental, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental de la empresa.

### 3.3 Regresión lineal simple con Responsabilidad Ambiental

Se procede a realizar una regresión lineal simple con la variable RAmbiental como variable dependiente y la variable RSE como variable independiente, para ello se utiliza la función lm, la cual permite realizar una regresión lineal simple.

```
# Regresion lineal simple con RAmbiental como variable dependiente

regresion_lineal_simple <- lm(RAmbiental ~ RSE, data = Percepcion_de_la_RSE)

summary(regresion_lineal_simple)

##
## Call:
## lm(formula = RAmbiental ~ RSE, data = Percepcion_de_la_RSE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.0088 -0.3580  0.1226  0.5059  1.6174
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.57159     0.37502   1.524    0.13
## RSE          0.77674     0.07151  10.862 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8587 on 108 degrees of freedom
## Multiple R-squared:  0.5221, Adjusted R-squared:  0.5177
## F-statistic: 118 on 1 and 108 DF, p-value: < 2.2e-16

#nolint
```

#### 3.3.1 Interpretación de la regresión lineal simple con Responsabilidad Ambiental como variable dependiente

Se observa que el modelo de regresión lineal simple con la variable RAmbiental como variable dependiente y la variable RSE como variable independiente, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de la variable RAmbiental. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental de la empresa.

#### 3.4 Interpretación del modelo de regresión en $R^2$

Se procede a interpretar el modelo de regresión en  $R^2$ , para ello se utiliza la función lm.beta, la cual permite obtener los coeficientes estandarizados del modelo de regresión.

Los coeficientes son significativos al 95% si el valor p ( $\Pr(>|t|)$ ) es menor a 0.05.

```
# Interpretacion del modelo de regresion en R^2
```

```
#library(car)
```

```
#std.coef(regresion_lineal_simple)
```

```
#nolint
```

### 3.4.1 Interpretación del modelo de regresión en $R^2$

Se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental de la empresa.

### 3.5 Modelo de regresión lineal múltiple

Se procede a realizar un modelo de regresión lineal múltiple con las variables RAmbiental y REconómica como variables dependientes y la variable RSE como variable independiente, para ello se utiliza la función lm, la cual permite realizar un modelo de regresión lineal múltiple.

```
# Modelo de regresion lineal multiple con RAmbiental y REconómica como variables dependientes y RSE como
```

```
regresion_lineal_multiple <- lm(cbind(RAmbiental, REconómica) ~ RSE, data = Percepcion_de_la_RSE) #nolint
```

```
summary(regresion_lineal_multiple) #nolint
```

```
## Response RAmbiental :
```

```
##
```

```
## Call:
```

```
## lm(formula = RAmbiental ~ RSE, data = Percepcion_de_la_RSE)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -5.0088 -0.3580  0.1226  0.5059  1.6174
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  0.57159    0.37502   1.524    0.13
```

```
## RSE          0.77674    0.07151  10.862 <2e-16 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 0.8587 on 108 degrees of freedom
```

```
## Multiple R-squared:  0.5221, Adjusted R-squared:  0.5177
```

```
## F-statistic: 118 on 1 and 108 DF, p-value: < 2.2e-16
```

```
##
```

```
##
```

```
## Response REconómica :
```

```
##
```

```
## Call:
```

```
## lm(formula = REconómica ~ RSE, data = Percepcion_de_la_RSE)
```

```
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.14145 -0.38011  0.09807  0.34709  2.20210
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.70458    0.27453  13.494 < 2e-16 ***
## RSE           0.47224    0.05235   9.022 7.76e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6286 on 108 degrees of freedom
## Multiple R-squared:  0.4298, Adjusted R-squared:  0.4245
## F-statistic: 81.39 on 1 and 108 DF,  p-value: 7.762e-15
```

```
#nolint
```

### 3.5.1 Interpretación del modelo de regresión lineal múltiple

Se observa que el modelo de regresión lineal múltiple con las variables RAmbiental y REconómica como variables dependientes y la variable RSE como variable independiente, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de las variables RAmbiental y REconómica. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental y Responsabilidad Económica de la empresa.

Elegimos la variable REconómica como variable dependiente, ya que presenta un coeficiente estandarizado mayor que la variable RAmbiental.

### 3.6 Regresión lineal múltiple con Responsabilidad Económica

Se procede a realizar una regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, para ello se utiliza la función `lm`, la cual permite realizar una regresión lineal múltiple.

```
# Regresion lineal multiple con REconómica como variable dependiente
```

```
regresion_lineal_multiple <- lm(REconómica ~ RSE + SaludSeguridad, data = Percepcion_de_la_RSE) #nolint
summary(regresion_lineal_multiple)
```

```
##
## Call:
## lm(formula = REconómica ~ RSE + SaludSeguridad, data = Percepcion_de_la_RSE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.13931 -0.37902  0.09912  0.38937  2.18180
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.61613    0.31541  11.465 < 2e-16 ***
```

```
## RSE          0.44646    0.06904    6.467 3.07e-09 ***
## SaludSeguridad 0.04198    0.07298    0.575    0.566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6306 on 107 degrees of freedom
## Multiple R-squared:  0.4315, Adjusted R-squared:  0.4209
## F-statistic: 40.61 on 2 and 107 DF,  p-value: 7.546e-14
```

```
#nolint
```

### 3.6.1 Interpretación de la regresión lineal múltiple con Responsabilidad Económica como variable dependiente

Se observa que el modelo de regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de la variable REconómica. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Económica de la empresa. Por otra parte, se observa que el coeficiente de la variable SaludSeguridad es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la Salud y Seguridad de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Económica de la empresa.

### 3.6.2 Test de significancia global del modelo de regresión lineal múltiple con Responsabilidad Económica como variable dependiente

Se procede a realizar un test de significancia global del modelo de regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, para ello se utiliza la función anova, la cual permite realizar un test de significancia global.

```
# Test de significancia global del modelo de regresion lineal multiple con REconómica como variable dep
```

```
anova(regresion_lineal_multiple)
```

```
## Analysis of Variance Table
##
## Response: REconómica
##          Df Sum Sq Mean Sq F value    Pr(>F)
## RSE          1 32.161   32.161 80.8867 9.61e-15 ***
## SaludSeguridad 1  0.132    0.132  0.3308  0.5664
## Residuals    107 42.543    0.398
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#nolint
```

### 3.6.3 Interpretación del test de significancia global del modelo de regresión lineal múltiple con Responsabilidad Económica como variable dependiente

Se observa que el test de significancia global del modelo de regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, presenta un valor p menor a 0.05, lo cual indica que el modelo es significativo.

#### 4. Conclusiones

- Se observa que la variable RSE presenta una correlación positiva y significativa con las variables RAmbiental, REconómica, RSocial y SaludSeguridad, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental, Responsabilidad Económica, Responsabilidad Social y Salud y Seguridad de la empresa.
- Se observa que el modelo de regresión lineal simple con la variable RAmbiental como variable dependiente y la variable RSE como variable independiente, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de la variable RAmbiental. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental de la empresa.
- Se observa que el modelo de regresión lineal múltiple con las variables RAmbiental y REconómica como variables dependientes y la variable RSE como variable independiente, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de las variables RAmbiental y REconómica. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Ambiental y Responsabilidad Económica de la empresa.
- Se observa que el modelo de regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, presenta un R cuadrado ajustado de 0.72, lo cual indica que el modelo explica el 72% de la variabilidad de la variable REconómica. Por otra parte, se observa que el coeficiente de la variable RSE es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la RSE de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Económica de la empresa. Por otra parte, se observa que el coeficiente de la variable SaludSeguridad es positivo y significativo, lo cual indica que a mayor percepción de los trabajadores respecto a la Salud y Seguridad de la empresa, mayor será la percepción de los trabajadores respecto a la Responsabilidad Económica de la empresa.
- Se observa que el test de significancia global del modelo de regresión lineal múltiple con la variable REconómica como variable dependiente y las variables RSE y SaludSeguridad como variables independientes, presenta un valor p menor a 0.05, lo cual indica que el modelo es significativo.

#### 5. Recomendaciones

- Se recomienda a la empresa SERVSAT, enfocar sus esfuerzos en mejorar las variables que inciden negativamente en la percepción de los trabajadores respecto a la RSE de la empresa, con el fin de mejorar la percepción de los trabajadores respecto a la RSE de la empresa.

#### 6. Referencias

- <https://www.rdocumentation.org/packages/QuantPsyc/versions/1.5.4/topics/lm.beta>
- <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/anova>
- <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor>
- <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/lm>
- <https://www.rdocumentation.org/packages/ggplot2/versions/3.3.2/topics/ggplot>
- <https://www.rdocumentation.org/packages/foreign/versions/0.8-81/topics/read.spss>
- <https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/summary>
- <https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/view>