

Estadística para la Ciencia de Datos

Exámen final: Estadística y probabilidad aplicada

Manuel Alejandro Serrano Macias

Fecha de modificación: 2025-11-07 06:50:18.978556

Contents

| | |
|---|-----------|
| Reporte de Análisis Estadístico Exploratorio de Marginación Municipal 2020 | 1 |
| Prerequisitos | 1 |
| a) Rango, rango intercuantílico, media, mediana y desviación estándar | 1 |
| b) Gráfico de correlación | 13 |
| c) Gráfico de densidad del Índice de marginación normalizado | 14 |
| d) Población Total Índice de Marginación | 16 |
| e) Comparación Índice de Marginación | 18 |
| Distribuciones de Probabilidad | 19 |
| 1. Ejercicio concurso de TV | 19 |
| 2. Ejercicio células de levadura | 20 |
| 3. Ejercicio tornillos | 21 |
| 4. Ejercicio Computadoras | 23 |

Reporte de Análisis Estadístico Exploratorio de Marginación Municipal 2020

A continuación se presenta un análisis estadístico exploratorio del Índice de Rezago Municipal 2020. Los datos son extraídos del Consejo Nacional de Población (CONAPO)

Prerequisitos

1. Se instalan y preparan las librerías a utilizar durante el análisis.

```
#install.packages("readxl")
#install.packages("corrplot")

library(readxl)
library(corrplot)
```

2. Carga y procesamiento de datos para el análisis. Se cargan los datos de Marginación Municipal 2020 (CONAPO)

```
marginacion_base <- read_excel("IMM_2020.xlsx", sheet = "IMM_2020")
```

a) Rango, rango intercuantílico, media, mediana y desviación estándar

Rango, rango intercuantílico, media, mediana y desviación estándar de las nueve variables porcentuales (ANALF, SBASC, OVSDE, OVSEE, OVSAE, OVPT, VHAC, PL.5000, PO2SM) de cada una de las 32

entidades federativas.

Se definen las variables que incluiremos en el estudio, iteramos en cada estado para calcular las estadísticas de interés de cada variable mediante las siguientes instrucciones:

```
variables_estudio <-  
  c("ANALF", "SBASC", "OVSD", "OVSEE", "OVSAE", "OVPT", "VHAC", "PL.5000", "PO2SM")  
  
entidades <- unique(marginacion_base$NOM_ENT)  
  
for (variable in variables_estudio) {  
  marginacion_base[[variable]] <- as.numeric(marginacion_base[[variable]])  
}  
  
resultados <- list()  
  
for (estado in entidades) {  
  datos_estado <- marginacion_base[marginacion_base$NOM_ENT == estado, ]  
  estadisticas_estado <- list()  
  
  for (variable in variables_estudio) {  
    estado_actual <- datos_estado[[variable]]  
    rango <- range(estado_actual, na.rm = TRUE)  
    rango_intercuantilico <- IQR(estado_actual, na.rm = TRUE)  
    media <- mean(estado_actual, na.rm = TRUE)  
    mediana <- median(estado_actual, na.rm = TRUE)  
    desviacion_estandar <- sd(estado_actual, na.rm = TRUE)  
  
    calculo_estadisticas <- data.frame(  
      Variable = variable,  
      Rango_Min = rango[1],  
      Rango_Max = rango[2],  
      Rango_Intercuantilico = rango_intercuantilico,  
      Media = media,  
      Mediana = mediana,  
      Desviacion_Estandar = desviacion_estandar  
    )  
  
    estadisticas_estado[[variable]] <- calculo_estadisticas  
  }  
  
  df_estado <- do.call(rbind, estadisticas_estado)  
  rownames(df_estado) <- NULL  
  
  resultados[[estado]] <- df_estado  
}
```

Finalmente, se crea una lista principal de data frames para cada estado.

NOTA: Originalmente hacia el proceso con un ciclo, sin embargo, las tablas no es el formato que específico en el encabezado del Markdown para impresión de se imprimían en el pdf de manera amigable, investigué y se debe a que *kable*, que data frames no puede imprimir de manera adecuada dentro de un ciclo, la documentación recomienda tener un chunk de R por cada tabla para así tenerla impresa en el PDF de manera amigable. Seguramente hay una mejor manera de hacerlo, lo investigaré para futuros reportes. En este ejercicio se imprimió cada estado individualmente.

```
# Código original solo para referencia

# for (estado in sort(names(resultados))) {
#   dataframe_estado <- resultados[[estado]]
#   #
#   cols_numericas <- c("Rango_Min", "Rango_Max", "Rango_Intercuantilico",
#                       "Media", "Mediana", "Desviacion_Estandar")
#   #
#   dataframe_estado[cols_numericas] <-
#     lapply(dataframe_estado[cols_numericas], round, digits = 3)
#   #
#   print(knitr::kable(dataframe_estado, caption = estado))
# }

imprimir_resultados <- function(data_frame, estado) {
  knitr::kable(data_frame,
               caption = estado,
               digits = 3,
               booktabs = TRUE,
               longtable = FALSE)
}

imprimir_resultados(resultados[["Aguascalientes"]], "Aguascalientes")
```

Table 1: Aguascalientes

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.645 | 4.492 | 1.302 | 3.118 | 3.145 | 0.890 |
| SBASC | 20.367 | 42.482 | 7.439 | 30.063 | 28.420 | 6.001 |
| OVSDE | 0.105 | 2.650 | 1.425 | 1.056 | 0.713 | 0.922 |
| OVSEE | 0.113 | 1.031 | 0.191 | 0.530 | 0.501 | 0.259 |
| OVSAB | 0.379 | 1.781 | 0.335 | 0.969 | 0.860 | 0.446 |
| OVPT | 0.591 | 1.453 | 0.535 | 1.074 | 1.040 | 0.314 |
| VHAC | 10.340 | 22.986 | 3.475 | 19.687 | 20.976 | 3.749 |
| PL.5000 | 7.524 | 78.221 | 28.471 | 49.606 | 44.285 | 21.437 |
| PO2SM | 54.227 | 81.726 | 12.651 | 71.461 | 76.336 | 9.531 |

```
imprimir_resultados(resultados[["Baja California"]], "Baja California")
```

Table 2: Baja California

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.470 | 8.122 | 0.522 | 2.913 | 1.990 | 2.569 |
| SBASC | 21.501 | 47.242 | 4.378 | 29.268 | 26.270 | 9.227 |
| OVSDE | 0.155 | 0.523 | 0.149 | 0.305 | 0.325 | 0.138 |
| OVSEE | 0.184 | 5.494 | 0.893 | 1.484 | 0.709 | 2.015 |
| OVSAB | 0.676 | 16.808 | 3.223 | 5.244 | 3.905 | 5.930 |
| OVPT | 1.226 | 4.437 | 0.896 | 2.308 | 1.932 | 1.155 |
| VHAC | 12.588 | 31.074 | 1.440 | 16.978 | 14.494 | 6.975 |
| PL.5000 | 1.704 | 54.398 | 7.304 | 19.538 | 14.663 | 18.167 |
| PO2SM | 65.782 | 79.432 | 3.835 | 72.750 | 73.399 | 4.612 |

```
imprimir_resultados(resultados[["Baja California Sur"]], "Baja California Sur")
```

Table 3: Baja California Sur

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.783 | 4.153 | 1.962 | 2.871 | 2.212 | 1.150 |
| SBASC | 19.889 | 33.589 | 8.386 | 27.157 | 26.032 | 5.756 |
| OVSDE | 0.243 | 1.037 | 0.584 | 0.651 | 0.521 | 0.359 |
| OVSEE | 0.585 | 2.177 | 0.497 | 1.185 | 1.077 | 0.616 |
| OVSDE | 1.552 | 9.289 | 0.560 | 3.799 | 2.857 | 3.118 |
| OVPT | 2.106 | 8.352 | 0.819 | 4.087 | 3.277 | 2.459 |
| VHAC | 11.965 | 24.588 | 5.089 | 18.653 | 19.351 | 4.832 |
| PL.5000 | 3.455 | 39.318 | 9.995 | 16.375 | 9.819 | 14.073 |
| PO2SM | 38.139 | 62.555 | 8.356 | 52.213 | 55.326 | 9.352 |

```
imprimir_resultados(resultados[["Campeche"]], "Campeche")
```

Table 4: Campeche

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 3.309 | 13.200 | 1.575 | 8.191 | 8.595 | 2.747 |
| SBASC | 22.971 | 44.089 | 7.385 | 34.837 | 34.423 | 6.334 |
| OVSDE | 0.667 | 8.649 | 3.797 | 4.406 | 3.751 | 2.700 |
| OVSEE | 0.341 | 6.980 | 0.805 | 1.530 | 1.091 | 1.799 |
| OVSDE | 0.655 | 21.903 | 5.979 | 6.268 | 2.623 | 7.203 |
| OVPT | 0.825 | 9.936 | 2.809 | 3.487 | 2.512 | 2.988 |
| VHAC | 24.155 | 46.496 | 4.818 | 34.485 | 35.114 | 6.024 |
| PL.5000 | 9.925 | 100.000 | 50.285 | 49.137 | 42.720 | 29.523 |
| PO2SM | 55.029 | 89.484 | 5.186 | 79.987 | 81.915 | 9.641 |

```
imprimir_resultados(resultados[["Coahuila de Zaragoza"]], "Coahuila de Zaragoza")
```

Table 5: Coahuila de Zaragoza

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.186 | 7.420 | 1.826 | 2.871 | 2.850 | 1.452 |
| SBASC | 16.410 | 50.039 | 10.479 | 30.418 | 29.932 | 8.094 |
| OVSDE | 0.106 | 3.074 | 0.783 | 0.719 | 0.594 | 0.655 |
| OVSEE | 0.000 | 3.703 | 0.294 | 0.472 | 0.272 | 0.714 |
| OVSDE | 0.081 | 9.832 | 1.074 | 1.603 | 0.839 | 2.131 |
| OVPT | 0.000 | 2.212 | 0.685 | 0.766 | 0.587 | 0.540 |
| VHAC | 8.840 | 22.912 | 4.167 | 16.465 | 17.411 | 3.116 |
| PL.5000 | 0.329 | 100.000 | 92.641 | 46.995 | 31.733 | 43.092 |
| PO2SM | 46.297 | 93.680 | 17.867 | 70.690 | 73.119 | 12.407 |

```
imprimir_resultados(resultados[["Colima"]], "Colima")
```

Table 6: Colima

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.319 | 7.363 | 4.185 | 4.590 | 4.498 | 2.248 |
| SBASC | 16.655 | 46.372 | 13.217 | 34.022 | 35.169 | 9.439 |
| OVSDE | 0.069 | 1.306 | 0.698 | 0.598 | 0.580 | 0.434 |
| OVSEE | 0.160 | 1.241 | 0.649 | 0.570 | 0.433 | 0.377 |
| OVSAB | 0.159 | 2.223 | 0.836 | 1.090 | 1.175 | 0.662 |
| OVPT | 0.979 | 7.335 | 3.372 | 3.721 | 3.924 | 2.193 |
| VHAC | 10.228 | 22.908 | 4.630 | 17.793 | 19.352 | 4.299 |
| PL.5000 | 1.513 | 100.000 | 32.136 | 33.665 | 32.799 | 29.532 |
| PO2SM | 50.694 | 83.405 | 14.562 | 67.009 | 68.304 | 10.138 |

```
imprimir_resultados(resultados[["Chiapas"]], "Chiapas")
```

Table 7: Chiapas

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 3.769 | 40.108 | 9.561 | 16.403 | 14.229 | 7.019 |
| SBASC | 23.413 | 86.142 | 14.102 | 54.041 | 53.433 | 11.524 |
| OVSDE | 0.139 | 45.440 | 1.955 | 2.843 | 1.921 | 4.469 |
| OVSEE | 0.238 | 11.558 | 1.230 | 1.759 | 1.213 | 1.731 |
| OVSAB | 0.204 | 70.051 | 9.069 | 10.512 | 7.696 | 10.266 |
| OVPT | 1.956 | 41.198 | 9.552 | 13.543 | 10.825 | 7.923 |
| VHAC | 21.374 | 63.512 | 12.612 | 41.150 | 39.700 | 9.262 |
| PL.5000 | 1.597 | 100.000 | 41.856 | 75.257 | 77.650 | 25.484 |
| PO2SM | 69.091 | 99.321 | 7.831 | 91.051 | 92.177 | 6.032 |

```
imprimir_resultados(resultados[["Chihuahua"]], "Chihuahua")
```

Table 8: Chihuahua

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.017 | 45.523 | 4.620 | 7.481 | 3.817 | 8.490 |
| SBASC | 16.414 | 84.481 | 20.621 | 49.422 | 51.509 | 14.910 |
| OVSDE | 0.071 | 57.185 | 2.939 | 5.214 | 0.975 | 10.371 |
| OVSEE | 0.062 | 53.065 | 3.239 | 5.550 | 0.860 | 10.510 |
| OVSAB | 0.030 | 36.010 | 1.477 | 4.096 | 0.593 | 8.044 |
| OVPT | 0.101 | 56.070 | 2.341 | 6.097 | 0.915 | 11.827 |
| VHAC | 5.340 | 47.251 | 8.458 | 16.027 | 12.587 | 9.139 |
| PL.5000 | 0.721 | 100.000 | 58.731 | 74.874 | 100.000 | 36.777 |
| PO2SM | 42.117 | 93.464 | 15.647 | 73.748 | 74.600 | 11.610 |

```
imprimir_resultados(resultados[["Ciudad de México"]], "Ciudad de México")
```

Table 9: Ciudad de México

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|-------|---------|---------------------|
| ANALF | 0.353 | 2.770 | 0.646 | 1.422 | 1.517 | 0.570 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| SBASC | 5.535 | 23.306 | 5.593 | 17.150 | 18.379 | 4.633 |
| OVSDE | 0.007 | 0.455 | 0.031 | 0.073 | 0.048 | 0.107 |
| OVSEE | 0.010 | 0.289 | 0.048 | 0.067 | 0.029 | 0.079 |
| OVSAB | 0.020 | 10.589 | 1.466 | 1.856 | 0.226 | 3.321 |
| OVPT | 0.068 | 2.577 | 0.723 | 0.778 | 0.485 | 0.786 |
| VHAC | 3.950 | 25.895 | 6.442 | 14.668 | 15.218 | 5.178 |
| PL.5000 | 0.000 | 24.412 | 2.424 | 2.310 | 0.000 | 6.046 |
| PO2SM | 28.453 | 72.965 | 13.249 | 55.591 | 57.004 | 11.081 |

```
imprimir_resultados(resultados[["Durango"]], "Durango")
```

Table 10: Durango

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.403 | 22.869 | 2.300 | 4.651 | 3.519 | 3.567 |
| SBASC | 19.063 | 58.932 | 12.635 | 40.571 | 39.235 | 8.969 |
| OVSDE | 0.467 | 38.950 | 4.173 | 5.398 | 3.084 | 6.730 |
| OVSEE | 0.080 | 42.405 | 3.086 | 3.189 | 0.741 | 7.059 |
| OVSAB | 0.201 | 50.631 | 2.556 | 3.511 | 1.153 | 8.673 |
| OVPT | 0.955 | 50.820 | 3.483 | 6.212 | 1.907 | 10.270 |
| VHAC | 9.641 | 55.670 | 8.435 | 19.575 | 16.577 | 8.902 |
| PL.5000 | 10.546 | 100.000 | 49.600 | 72.938 | 68.767 | 29.148 |
| PO2SM | 60.674 | 93.166 | 9.552 | 79.073 | 80.810 | 8.427 |

```
imprimir_resultados(resultados[["Guanajuato"]], "Guanajuato")
```

Table 11: Guanajuato

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 3.313 | 14.619 | 3.170 | 7.638 | 7.663 | 2.566 |
| SBASC | 23.845 | 51.324 | 8.803 | 40.120 | 41.542 | 7.274 |
| OVSDE | 0.190 | 15.189 | 4.332 | 3.536 | 2.219 | 3.477 |
| OVSEE | 0.106 | 2.760 | 0.581 | 0.647 | 0.394 | 0.604 |
| OVSAB | 0.243 | 41.668 | 2.847 | 3.570 | 1.480 | 6.560 |
| OVPT | 0.566 | 6.763 | 2.368 | 2.668 | 2.301 | 1.572 |
| VHAC | 8.655 | 29.826 | 7.149 | 19.303 | 18.994 | 4.477 |
| PL.5000 | 7.009 | 100.000 | 32.845 | 58.435 | 54.958 | 26.103 |
| PO2SM | 57.462 | 85.865 | 11.739 | 76.036 | 79.370 | 8.042 |

```
imprimir_resultados(resultados[["Guerrero"]], "Guerrero")
```

Table 12: Guerrero

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 5.266 | 53.071 | 8.655 | 17.734 | 16.000 | 8.799 |
| SBASC | 25.462 | 82.819 | 17.217 | 53.523 | 52.122 | 11.837 |
| OVSDE | 0.796 | 61.040 | 13.172 | 13.455 | 9.844 | 11.437 |
| OVSEE | 0.240 | 14.128 | 1.792 | 2.201 | 1.284 | 2.417 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| OVSAB | 2.239 | 68.852 | 11.438 | 14.066 | 9.572 | 12.622 |
| OVPT | 3.214 | 59.684 | 13.374 | 19.005 | 14.753 | 11.872 |
| VHAC | 20.309 | 64.029 | 16.242 | 38.175 | 36.203 | 10.015 |
| PL.5000 | 11.375 | 100.000 | 48.491 | 74.068 | 79.291 | 28.072 |
| PO2SM | 68.674 | 97.670 | 9.681 | 87.109 | 88.271 | 6.819 |

```
imprimir_resultados(resultados[["Hidalgo"]], "Hidalgo")
```

Table 13: Hidalgo

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.225 | 26.744 | 9.646 | 9.344 | 7.315 | 6.652 |
| SBASC | 13.029 | 56.352 | 15.622 | 36.346 | 35.672 | 10.399 |
| OVSDE | 0.094 | 11.420 | 3.131 | 2.812 | 1.983 | 2.585 |
| OVSEE | 0.103 | 4.297 | 0.711 | 0.974 | 0.812 | 0.750 |
| OVSAB | 0.320 | 34.817 | 4.036 | 5.434 | 2.276 | 7.371 |
| OVPT | 0.415 | 17.855 | 2.898 | 3.643 | 2.369 | 3.477 |
| VHAC | 9.120 | 40.134 | 7.954 | 21.509 | 19.707 | 6.673 |
| PL.5000 | 5.244 | 100.000 | 49.974 | 74.302 | 100.000 | 30.444 |
| PO2SM | 54.529 | 97.764 | 13.733 | 81.419 | 82.942 | 9.509 |

```
imprimir_resultados(resultados[["Jalisco"]], "Jalisco")
```

Table 14: Jalisco

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.526 | 24.747 | 3.041 | 5.785 | 5.265 | 3.111 |
| SBASC | 19.637 | 63.111 | 12.505 | 43.932 | 45.583 | 9.138 |
| OVSDE | 0.010 | 60.416 | 1.395 | 2.346 | 0.738 | 6.692 |
| OVSEE | 0.017 | 30.727 | 0.637 | 1.129 | 0.324 | 3.370 |
| OVSAB | 0.031 | 31.013 | 0.996 | 1.690 | 0.820 | 3.397 |
| OVPT | 0.242 | 46.077 | 2.304 | 2.939 | 1.482 | 5.247 |
| VHAC | 8.670 | 41.584 | 6.706 | 18.049 | 17.497 | 5.398 |
| PL.5000 | 0.001 | 100.000 | 73.997 | 55.770 | 42.918 | 35.259 |
| PO2SM | 46.149 | 88.566 | 9.171 | 68.844 | 69.072 | 7.814 |

```
imprimir_resultados(resultados[["México"]], "México")
```

Table 15: México

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 0.736 | 14.870 | 4.496 | 4.774 | 3.328 | 3.421 |
| SBASC | 13.821 | 57.017 | 16.428 | 31.315 | 29.218 | 10.646 |
| OVSDE | 0.012 | 20.250 | 5.013 | 3.140 | 0.775 | 4.404 |
| OVSEE | 0.020 | 2.685 | 0.533 | 0.567 | 0.332 | 0.597 |
| OVSAB | 0.000 | 30.338 | 4.704 | 4.035 | 2.279 | 4.653 |
| OVPT | 0.330 | 11.002 | 2.725 | 3.147 | 2.491 | 2.167 |
| VHAC | 6.586 | 41.956 | 6.999 | 24.073 | 24.078 | 6.029 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| PL.5000 | 0.000 | 100.000 | 67.937 | 49.258 | 47.642 | 36.043 |
| PO2SM | 44.794 | 95.947 | 12.474 | 73.246 | 73.293 | 8.834 |

```
imprimir_resultados(resultados[["Michoacán de Ocampo"]], "Michoacán de Ocampo")
```

Table 16: Michoacán de Ocampo

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 2.800 | 18.437 | 3.794 | 9.608 | 9.149 | 3.263 |
| SBASC | 23.977 | 73.643 | 8.801 | 52.286 | 52.819 | 7.914 |
| OVSDE | 0.069 | 18.376 | 2.150 | 2.485 | 1.175 | 3.297 |
| OVSEE | 0.084 | 10.665 | 0.483 | 0.800 | 0.402 | 1.291 |
| OVSAE | 0.211 | 25.915 | 2.226 | 3.044 | 1.945 | 3.655 |
| OVPT | 0.157 | 40.483 | 6.542 | 6.646 | 5.114 | 6.213 |
| VHAC | 7.816 | 36.496 | 8.430 | 21.490 | 21.204 | 5.742 |
| PL.5000 | 8.964 | 100.000 | 63.267 | 62.945 | 61.721 | 31.633 |
| PO2SM | 51.762 | 93.620 | 10.699 | 77.659 | 79.711 | 8.647 |

```
imprimir_resultados(resultados[["Morelos"]], "Morelos")
```

Table 17: Morelos

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 2.199 | 10.612 | 2.189 | 5.639 | 5.318 | 1.863 |
| SBASC | 18.960 | 57.056 | 6.439 | 33.329 | 33.449 | 7.615 |
| OVSDE | 0.113 | 3.210 | 1.262 | 1.373 | 1.297 | 0.829 |
| OVSEE | 0.077 | 0.955 | 0.293 | 0.411 | 0.371 | 0.216 |
| OVSAE | 0.445 | 34.199 | 5.752 | 6.709 | 2.810 | 8.251 |
| OVPT | 1.236 | 27.144 | 2.886 | 5.776 | 4.490 | 4.999 |
| VHAC | 13.605 | 34.573 | 5.788 | 22.666 | 22.441 | 4.714 |
| PL.5000 | 4.662 | 100.000 | 40.348 | 44.327 | 46.173 | 27.880 |
| PO2SM | 58.724 | 92.737 | 8.324 | 81.378 | 82.577 | 7.436 |

```
imprimir_resultados(resultados[["Nayarit"]], "Nayarit")
```

Table 18: Nayarit

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 2.131 | 27.268 | 1.848 | 6.744 | 5.542 | 5.363 |
| SBASC | 19.298 | 66.476 | 9.207 | 36.763 | 36.225 | 10.801 |
| OVSDE | 0.212 | 64.450 | 3.430 | 7.486 | 1.300 | 15.514 |
| OVSEE | 0.133 | 32.724 | 1.027 | 4.125 | 0.541 | 8.723 |
| OVSAE | 0.389 | 31.887 | 2.472 | 4.821 | 1.904 | 7.845 |
| OVPT | 0.699 | 47.535 | 3.148 | 6.848 | 1.890 | 12.118 |
| VHAC | 12.291 | 54.016 | 8.611 | 22.132 | 19.894 | 10.125 |
| PL.5000 | 10.995 | 100.000 | 63.991 | 61.383 | 66.538 | 32.623 |
| PO2SM | 55.006 | 87.823 | 10.061 | 73.967 | 76.029 | 10.166 |


```
imprimir_resultados(resultados[["Nuevo León"]], "Nuevo León")
```

Table 19: Nuevo León

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 0.736 | 11.626 | 2.741 | 3.419 | 3.001 | 2.390 |
| SBASC | 11.660 | 57.504 | 21.973 | 33.986 | 35.140 | 12.840 |
| OVSDE | 0.000 | 4.950 | 0.706 | 0.742 | 0.331 | 1.068 |
| OVSEE | 0.000 | 2.438 | 0.881 | 0.630 | 0.310 | 0.691 |
| OVSAB | 0.021 | 29.325 | 1.971 | 3.101 | 1.408 | 5.724 |
| OVPT | 0.000 | 23.999 | 0.764 | 1.798 | 0.700 | 3.873 |
| VHAC | 6.203 | 30.516 | 5.866 | 17.061 | 17.214 | 5.408 |
| PL.5000 | 0.000 | 100.000 | 95.814 | 48.931 | 18.266 | 45.690 |
| PO2SM | 32.818 | 93.439 | 21.576 | 62.646 | 62.378 | 14.599 |

```
imprimir_resultados(resultados[["Oaxaca"]], "Oaxaca")
```

Table 20: Oaxaca

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.107 | 49.801 | 12.400 | 14.492 | 12.634 | 8.834 |
| SBASC | 12.892 | 88.328 | 15.412 | 55.187 | 56.738 | 12.869 |
| OVSDE | 0.000 | 21.677 | 1.951 | 2.098 | 1.329 | 2.495 |
| OVSEE | 0.000 | 16.262 | 1.844 | 1.970 | 1.406 | 2.028 |
| OVSAB | 0.000 | 67.433 | 8.397 | 7.698 | 3.664 | 10.815 |
| OVPT | 0.458 | 68.150 | 12.031 | 15.054 | 13.105 | 10.172 |
| VHAC | 8.597 | 65.604 | 12.820 | 30.124 | 29.495 | 10.052 |
| PL.5000 | 0.000 | 100.000 | 0.000 | 90.504 | 100.000 | 24.870 |
| PO2SM | 46.729 | 100.000 | 9.109 | 89.039 | 91.863 | 8.910 |

```
imprimir_resultados(resultados[["Puebla"]], "Puebla")
```

Table 21: Puebla

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.447 | 32.784 | 8.575 | 12.123 | 10.688 | 7.014 |
| SBASC | 18.015 | 72.083 | 14.465 | 50.206 | 50.904 | 10.717 |
| OVSDE | 0.000 | 15.691 | 2.175 | 2.365 | 1.569 | 2.452 |
| OVSEE | 0.000 | 5.825 | 0.996 | 1.067 | 0.763 | 0.940 |
| OVSAB | 0.000 | 62.404 | 6.303 | 6.379 | 3.548 | 8.151 |
| OVPT | 0.103 | 33.629 | 6.419 | 7.992 | 6.347 | 6.038 |
| VHAC | 11.423 | 51.432 | 8.547 | 31.827 | 31.599 | 7.210 |
| PL.5000 | 1.284 | 100.000 | 57.715 | 72.915 | 100.000 | 33.940 |
| PO2SM | 58.410 | 99.091 | 7.247 | 89.416 | 91.489 | 6.879 |

```
imprimir_resultados(resultados[["Querétaro"]], "Querétaro")
```

Table 22: Querétaro

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.676 | 13.082 | 5.907 | 7.362 | 7.786 | 3.650 |
| SBASC | 14.257 | 50.626 | 15.344 | 37.436 | 41.527 | 11.367 |
| OVSDE | 0.079 | 12.757 | 5.782 | 4.351 | 2.016 | 4.301 |
| OVSEE | 0.100 | 5.004 | 1.596 | 1.334 | 0.918 | 1.272 |
| OVSAB | 0.672 | 26.977 | 8.391 | 6.043 | 2.933 | 6.814 |
| OVPT | 1.144 | 6.325 | 1.231 | 3.039 | 2.884 | 1.507 |
| VHAC | 7.376 | 34.169 | 10.137 | 23.039 | 23.137 | 7.473 |
| PL.5000 | 7.140 | 100.000 | 59.200 | 66.144 | 77.506 | 32.735 |
| PO2SM | 41.263 | 85.104 | 8.248 | 70.455 | 74.103 | 11.133 |

```
imprimir_resultados(resultados[["Quintana Roo"]], "Quintana Roo")
```

Table 23: Quintana Roo

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.728 | 11.119 | 6.006 | 5.551 | 4.078 | 3.556 |
| SBASC | 17.731 | 41.667 | 12.675 | 28.941 | 29.674 | 7.879 |
| OVSDE | 0.064 | 10.988 | 4.894 | 3.291 | 0.760 | 4.077 |
| OVSEE | 0.343 | 2.791 | 1.541 | 1.685 | 2.218 | 0.965 |
| OVSAB | 0.455 | 6.018 | 1.379 | 2.305 | 1.858 | 1.476 |
| OVPT | 0.772 | 10.168 | 4.379 | 4.363 | 3.754 | 2.941 |
| VHAC | 22.350 | 46.379 | 19.223 | 35.123 | 39.377 | 9.736 |
| PL.5000 | 0.320 | 72.113 | 57.964 | 30.992 | 25.281 | 29.658 |
| PO2SM | 51.265 | 85.594 | 25.070 | 67.309 | 66.551 | 13.273 |

```
imprimir_resultados(resultados[["San Luis Potosí"]], "San Luis Potosí")
```

Table 24: San Luis Potosí

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.763 | 21.717 | 3.816 | 8.685 | 8.755 | 3.418 |
| SBASC | 18.801 | 60.747 | 11.526 | 41.251 | 40.871 | 8.922 |
| OVSDE | 0.105 | 15.569 | 2.596 | 2.833 | 1.860 | 2.839 |
| OVSEE | 0.221 | 6.998 | 1.897 | 2.156 | 1.658 | 1.470 |
| OVSAB | 0.793 | 50.248 | 13.237 | 14.091 | 9.954 | 11.969 |
| OVPT | 0.963 | 26.488 | 7.716 | 7.844 | 6.210 | 5.733 |
| VHAC | 8.996 | 38.398 | 7.526 | 22.819 | 22.183 | 6.164 |
| PL.5000 | 3.686 | 100.000 | 51.842 | 71.597 | 76.284 | 31.458 |
| PO2SM | 47.749 | 95.738 | 11.487 | 83.639 | 86.633 | 9.357 |

```
imprimir_resultados(resultados[["Sinaloa"]], "Sinaloa")
```

Table 25: Sinaloa

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|-------|---------|---------------------|
| ANALF | 1.865 | 11.588 | 3.500 | 6.055 | 5.663 | 2.833 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| SBASC | 23.661 | 54.220 | 12.991 | 38.467 | 38.673 | 9.781 |
| OVSDE | 0.418 | 10.873 | 3.465 | 3.457 | 2.967 | 2.807 |
| OVSEE | 0.093 | 3.462 | 0.825 | 0.970 | 0.517 | 0.977 |
| OVSAB | 0.538 | 6.814 | 2.940 | 2.695 | 2.031 | 1.971 |
| OVPT | 0.978 | 15.483 | 3.262 | 4.285 | 3.792 | 3.504 |
| VHAC | 15.605 | 28.121 | 6.873 | 22.647 | 24.481 | 4.144 |
| PL.5000 | 6.659 | 100.000 | 34.455 | 55.153 | 62.355 | 28.176 |
| PO2SM | 53.321 | 83.791 | 8.913 | 71.699 | 72.137 | 8.025 |

```
imprimir_resultados(resultados[["Sonora"]], "Sonora")
```

Table 26: Sonora

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 0.887 | 10.258 | 1.691 | 3.198 | 2.784 | 1.953 |
| SBASC | 16.492 | 58.123 | 16.563 | 37.151 | 36.782 | 10.469 |
| OVSDE | 0.000 | 14.169 | 1.043 | 1.267 | 0.801 | 1.955 |
| OVSEE | 0.000 | 9.384 | 1.138 | 1.318 | 0.877 | 1.501 |
| OVSAB | 0.000 | 10.102 | 1.122 | 1.204 | 0.731 | 1.669 |
| OVPT | 0.000 | 12.351 | 1.883 | 2.259 | 1.357 | 2.680 |
| VHAC | 5.848 | 41.847 | 7.294 | 15.430 | 13.995 | 6.920 |
| PL.5000 | 0.619 | 100.000 | 74.073 | 71.489 | 100.000 | 39.385 |
| PO2SM | 43.014 | 93.447 | 15.361 | 71.075 | 70.467 | 10.847 |

```
imprimir_resultados(resultados[["Tabasco"]], "Tabasco")
```

Table 27: Tabasco

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 3.022 | 8.559 | 2.027 | 6.031 | 5.587 | 1.626 |
| SBASC | 21.588 | 39.274 | 8.051 | 32.070 | 33.101 | 5.489 |
| OVSDE | 0.421 | 5.115 | 1.516 | 2.374 | 2.536 | 1.359 |
| OVSEE | 0.136 | 1.941 | 0.601 | 0.702 | 0.517 | 0.550 |
| OVSAB | 0.927 | 22.693 | 4.516 | 6.872 | 5.104 | 6.040 |
| OVPT | 1.325 | 10.751 | 2.471 | 3.919 | 3.577 | 2.273 |
| VHAC | 21.961 | 35.910 | 5.637 | 28.313 | 26.857 | 4.166 |
| PL.5000 | 24.691 | 86.752 | 22.342 | 62.149 | 66.486 | 17.595 |
| PO2SM | 63.380 | 89.192 | 10.881 | 77.275 | 79.277 | 7.923 |

```
imprimir_resultados(resultados[["Tamaulipas"]], "Tamaulipas")
```

Table 28: Tamaulipas

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.357 | 15.373 | 3.261 | 5.408 | 5.006 | 2.968 |
| SBASC | 17.262 | 55.389 | 12.253 | 38.906 | 40.611 | 10.142 |
| OVSDE | 0.062 | 3.381 | 0.935 | 0.931 | 0.826 | 0.787 |
| OVSEE | 0.060 | 13.304 | 1.617 | 1.609 | 0.811 | 2.177 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| OVSAE | 0.172 | 81.788 | 5.306 | 7.504 | 3.326 | 14.294 |
| OVPT | 0.328 | 7.797 | 2.453 | 2.616 | 2.458 | 1.892 |
| VHAC | 7.015 | 34.296 | 5.627 | 19.465 | 19.160 | 5.653 |
| PL.5000 | 0.000 | 100.000 | 75.943 | 59.783 | 59.385 | 38.905 |
| PO2SM | 50.716 | 95.540 | 8.225 | 83.663 | 85.607 | 9.343 |

```
imprimir_resultados(resultados[["Tlaxcala"]], "Tlaxcala")
```

Table 29: Tlaxcala

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.239 | 9.618 | 2.240 | 3.567 | 2.956 | 1.849 |
| SBASC | 13.161 | 46.709 | 10.336 | 28.341 | 27.115 | 7.204 |
| OVSDE | 0.157 | 4.810 | 0.846 | 1.028 | 0.761 | 0.961 |
| OVSEE | 0.130 | 1.468 | 0.321 | 0.450 | 0.419 | 0.262 |
| OVSAE | 0.048 | 7.296 | 0.583 | 0.866 | 0.570 | 1.098 |
| OVPT | 0.123 | 10.684 | 1.390 | 1.814 | 1.225 | 1.737 |
| VHAC | 11.160 | 41.031 | 8.132 | 23.561 | 22.886 | 6.021 |
| PL.5000 | 0.000 | 100.000 | 98.850 | 42.339 | 27.692 | 40.049 |
| PO2SM | 61.426 | 96.703 | 7.867 | 81.713 | 82.866 | 6.263 |

```
imprimir_resultados(resultados[["Veracruz de Ignacio de la Llave"]], "Veracruz de Ignacio de la Llave")
```

Table 30: Veracruz de Ignacio de la Llave

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 2.123 | 41.401 | 7.608 | 12.523 | 11.002 | 6.998 |
| SBASC | 20.844 | 77.861 | 13.380 | 49.720 | 51.037 | 11.212 |
| OVSDE | 0.034 | 14.274 | 1.911 | 2.070 | 1.426 | 2.073 |
| OVSEE | 0.047 | 9.205 | 1.061 | 1.455 | 1.154 | 1.217 |
| OVSAE | 0.126 | 62.627 | 12.494 | 10.505 | 6.109 | 11.714 |
| OVPT | 0.233 | 48.412 | 6.680 | 7.711 | 5.625 | 7.134 |
| VHAC | 12.015 | 69.564 | 11.629 | 27.907 | 25.243 | 9.918 |
| PL.5000 | 0.000 | 100.000 | 56.611 | 69.887 | 75.069 | 31.967 |
| PO2SM | 59.110 | 98.764 | 9.778 | 86.557 | 88.707 | 8.545 |

```
imprimir_resultados(resultados[["Yucatán"]], "Yucatán")
```

Table 31: Yucatán

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 2.047 | 19.377 | 5.620 | 10.564 | 10.524 | 4.073 |
| SBASC | 19.529 | 61.561 | 11.209 | 44.351 | 44.872 | 8.012 |
| OVSDE | 0.697 | 40.759 | 9.115 | 11.438 | 9.899 | 6.887 |
| OVSEE | 0.142 | 5.301 | 1.042 | 1.078 | 0.875 | 0.909 |
| OVSAE | 0.058 | 5.790 | 1.218 | 1.392 | 1.066 | 1.182 |
| OVPT | 0.025 | 9.631 | 1.864 | 1.751 | 1.091 | 1.676 |
| VHAC | 16.001 | 61.104 | 9.649 | 35.380 | 35.262 | 8.112 |

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| PL.5000 | 0.033 | 100.000 | 72.832 | 67.147 | 100.000 | 40.524 |
| PO2SM | 55.836 | 94.808 | 6.999 | 85.997 | 87.261 | 6.549 |

```
imprimir_resultados(resultados[["Zacatecas"]], "Zacatecas")
```

Table 32: Zacatecas

| Variable | Rango_Min | Rango_Max | Rango_Intercuantilico | Media | Mediana | Desviacion_Estandar |
|----------|-----------|-----------|-----------------------|--------|---------|---------------------|
| ANALF | 1.442 | 13.698 | 1.712 | 5.094 | 5.006 | 1.924 |
| SBASC | 17.235 | 57.341 | 9.384 | 39.756 | 38.386 | 7.890 |
| OVSDE | 0.106 | 14.125 | 2.320 | 2.732 | 1.983 | 2.600 |
| OVSEE | 0.091 | 4.168 | 0.399 | 0.698 | 0.456 | 0.816 |
| OVSAB | 0.152 | 10.715 | 1.542 | 1.938 | 1.144 | 2.116 |
| OVPT | 0.286 | 8.065 | 0.665 | 1.488 | 1.232 | 1.167 |
| VHAC | 9.341 | 26.716 | 6.327 | 18.290 | 18.465 | 4.213 |
| PL.5000 | 3.438 | 100.000 | 56.385 | 70.234 | 73.424 | 31.819 |
| PO2SM | 56.511 | 90.877 | 9.482 | 78.839 | 80.537 | 8.146 |

b) Gráfico de correlación

Variables de estudio (ANALF, SBASC, OVSDE, OVSEE, OVSAB, OVPT, VHAC, PL.5000, PO2SM)

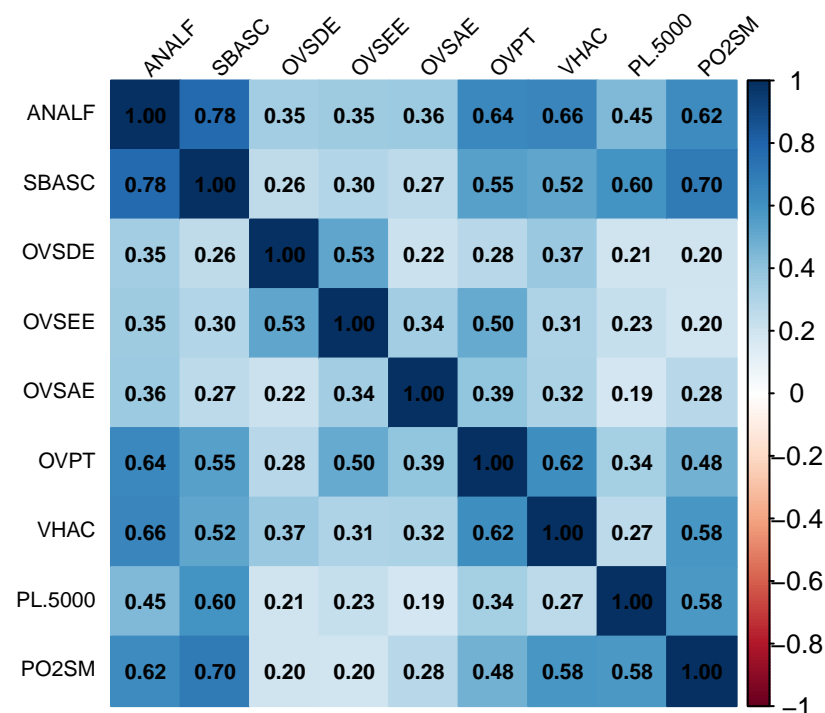
Para generar el gráfico de correlación creamos un nuevo data frame con las variables de estudio, posteriormente la función *cor* calcula la matriz de correlación de Pearson. Se utiliza la librería de *corrplot* para crear el gráfico.

```
datos_correlacion <- marginacion_base[, variables_estudio]
matriz_correlacion <- cor(datos_correlacion, use = "pairwise.complete.obs")
knitr::kable(round(matriz_correlacion, 2))
```

| | ANALF | SBASC | OVSDE | OVSEE | OVSAB | OVPT | VHAC | PL.5000 | PO2SM |
|---------|-------|-------|-------|-------|-------|------|------|---------|-------|
| ANALF | 1.00 | 0.78 | 0.35 | 0.35 | 0.36 | 0.64 | 0.66 | 0.45 | 0.62 |
| SBASC | 0.78 | 1.00 | 0.26 | 0.30 | 0.27 | 0.55 | 0.52 | 0.60 | 0.70 |
| OVSDE | 0.35 | 0.26 | 1.00 | 0.53 | 0.22 | 0.28 | 0.37 | 0.21 | 0.20 |
| OVSEE | 0.35 | 0.30 | 0.53 | 1.00 | 0.34 | 0.50 | 0.31 | 0.23 | 0.20 |
| OVSAB | 0.36 | 0.27 | 0.22 | 0.34 | 1.00 | 0.39 | 0.32 | 0.19 | 0.28 |
| OVPT | 0.64 | 0.55 | 0.28 | 0.50 | 0.39 | 1.00 | 0.62 | 0.34 | 0.48 |
| VHAC | 0.66 | 0.52 | 0.37 | 0.31 | 0.32 | 0.62 | 1.00 | 0.27 | 0.58 |
| PL.5000 | 0.45 | 0.60 | 0.21 | 0.23 | 0.19 | 0.34 | 0.27 | 1.00 | 0.58 |
| PO2SM | 0.62 | 0.70 | 0.20 | 0.20 | 0.28 | 0.48 | 0.58 | 0.58 | 1.00 |

```
corrplot(matriz_correlacion,
  method = "color",
  title = "Gráfico de Correlación de Indicadores de Marginación",
  mar = c(0, 0, 4, 0),
  cex.main = 1,
  tl.col = "black",
  tl.srt = 45,
  addCoef.col = "black",
  number.cex = 0.7,
  tl.cex = 0.7)
```

Gráfico de Correlación de Indicadores de Marginación



Observaciones

NOTA: La matriz de correlación es simétrica, aquí solo tendría la duda de si podría entonces graficarse solo la parte superior o inferior. Por el momento, la dejé completa.

1. Correlaciones Positivas Fuertes

- Educación y Pobreza de Ingresos: Se espera una correlación muy fuerte entre ANALF (analfabetismo) y SBASC (población sin educación básica) con PO2SM (población con ingresos ≤ 2 salarios mínimos). Esto significa que donde hay alta falta de educación básica, hay mayor población con bajos ingresos, reforzando el ciclo de marginación.
- Vivienda de Calidad: Los indicadores de carencias en la vivienda suelen estar muy correlacionados entre sí. OVSDE (sin drenaje/excusado), OVSEE (sin electricidad), OVSAE (sin agua entubada) y OVPT (piso de tierra) muestran una correlación positiva alta. Si un municipio carece de un servicio básico de vivienda, es muy probable que carezca de los demás.
- Hacinamiento (VHAC): También se correlaciona positivamente con las otras carencias de vivienda, pues ambas son manifestaciones de pobreza de infraestructura.

2. Correlaciones Débiles

- Correlación entre Desarrollo de Servicios y Variables Rurales: Es posible encontrar correlaciones más débiles entre variables como PL.5000 (población en localidades rurales/pequeñas) y algunos indicadores de vivienda. En general, en municipios con alta dispersión rural (PL.5000), las carencias de servicios (OVSDE, OVSAE) pueden ser elevadas.

Conclusión: La marginación es un fenómeno donde las carencias se acumulan mala vivienda, baja educación y bajos ingresos.

c) Gráfico de densidad del Índice de marginación normalizado

Estados de estudio: Nuevo León, Durango, Jalisco, Veracruz y Chiapas.

Para generar el gráfico de densidad se filtran los datos de las entidades federativas de interés, se calculan las

densidades de cada estado, posteriormente se determinan los rangos para que las curvas quepan y creamos el gráfico.

```
entidades_interes <- c("Nuevo León", "Durango", "Jalisco", "Veracruz de Ignacio de la Llave", "Chiapas")
variable_imn <- "IMN_2020"
datos_entidades_estudio <- marginacion_base[marginacion_base$NOM_ENT %in% entidades_interes, ]
datos_entidades_estudio[[variable_imn]] <- as.numeric(datos_entidades_estudio[[variable_imn]])

nuevo_leon_imn <- na.omit(subset(marginacion_base, NOM_ENT == "Nuevo León")[[variable_imn]])
durango_imn <- na.omit(subset(marginacion_base, NOM_ENT == "Durango")[[variable_imn]])
jalisco_imn <- na.omit(subset(marginacion_base, NOM_ENT == "Jalisco")[[variable_imn]])
veracruz_imn <- na.omit(subset(marginacion_base, NOM_ENT == "Veracruz de Ignacio de la Llave")[[variable_imn]])
chiapas_imn <- na.omit(subset(marginacion_base, NOM_ENT == "Chiapas")[[variable_imn]])

densidad_nuevo_leon <- density(nuevo_leon_imn)
densidad_durango <- density(durango_imn)
densidad_jalisco <- density(jalisco_imn)
densidad_veracruz <- density(veracruz_imn)
densidad_chiapas <- density(chiapas_imn)

colores <- c("cadetblue", "coral", "antiquewhite4", "brown", "goldenrod")

min_x_vals <- c(min(densidad_nuevo_leon$x),
                min(densidad_durango$x),
                min(densidad_jalisco$x),
                min(densidad_veracruz$x),
                min(densidad_chiapas$x))

max_x_vals <- c(max(densidad_nuevo_leon$x),
                max(densidad_durango$x),
                max(densidad_jalisco$x),
                max(densidad_veracruz$x),
                max(densidad_chiapas$x))

rango_x <- c(min(min_x_vals), max(max_x_vals))

max_y_vals <- c(max(densidad_nuevo_leon$y),
                max(densidad_durango$y),
                max(densidad_jalisco$y),
                max(densidad_veracruz$y),
                max(densidad_chiapas$y))

rango_y <- c(0, max(max_y_vals))

plot(densidad_nuevo_leon,
     xlim = rango_x,
     ylim = rango_y,
     main = "Densidad del Índice de Marginación Normalizado por Entidad",
     xlab = "Índice de Marginación Normalizado (IMN_2020)",
     ylab = "Densidad",
     col = colores[1],
     lwd = 2)

lines(densidad_durango, col = colores[2], lwd = 2)
```

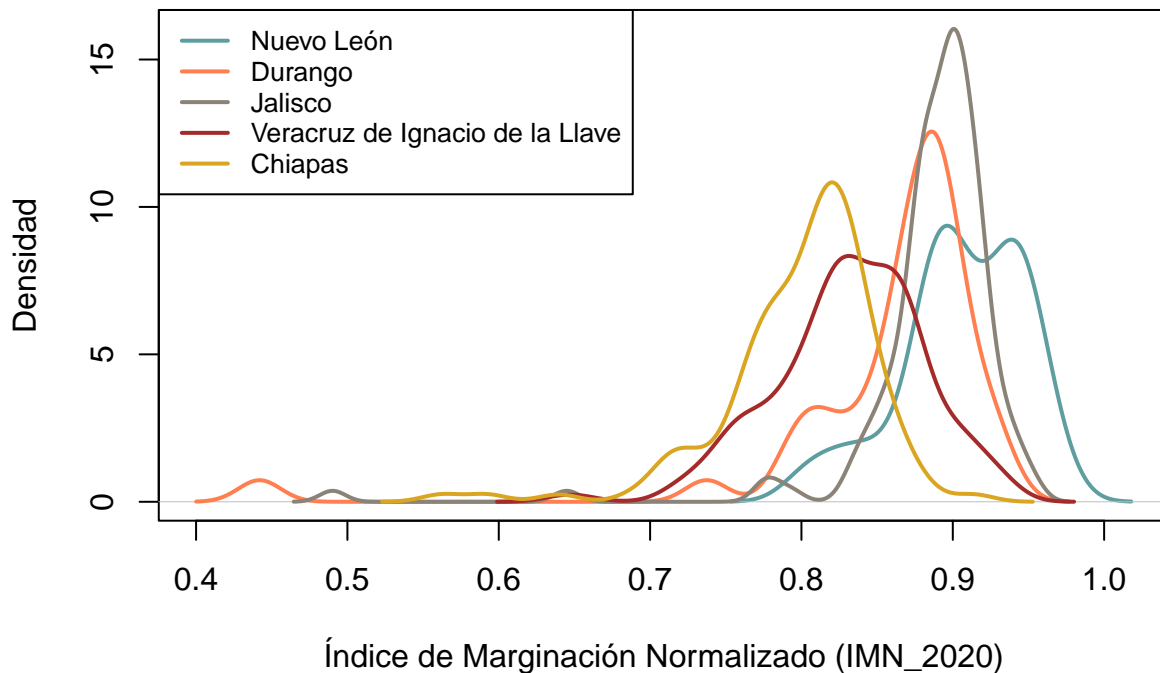
```

lines(densidad_jalisco, col = colores[3], lwd = 2)
lines(densidad_veracruz, col = colores[4], lwd = 2)
lines(densidad_chiapas, col = colores[5], lwd = 2)

legend("topleft",
      legend = entidades_interes,
      col = colores,
      lwd = 2,
      cex = 0.8)

```

Densidad del Índice de Marginación Normalizado por Entidad



Observaciones

- Nuevo León. La mayoría de sus municipios presentan una marginación muy baja.
- Jalisco. Coexisten municipios de muy baja marginación como la Zona Metropolitana de Guadalajara y un grupo considerable de municipios rurales con marginación media o alta.
- Durango. Indica una mezcla continua de municipios de baja y alta marginación.
- Veracruz. Presenta una alta dispersión de la marginación, la mayoría de sus municipios se encuentran en la zona de marginación media alta.
- Chiapas. Sus municipios tienen la marginación más alta y concentrada indicando que la mayor parte de la población municipal se ubica en los niveles más severos de marginación.

d) Población Total Índice de Marginación

A continuación se presenta el total de la población por cada variable del estudio.

Se calcula en base a la población total de los municipios de cada entidad federativa y el porcentaje de cada variable.

```

marginacion_base$POB_TOT <- as.numeric(marginacion_base$POB_TOT)

for (col in variables_estudio) {

```



```

marginacion_base[[col]] <- as.numeric(marginacion_base[[col]])
}

datos_filtrados <- marginacion_base[marginacion_base$NOM_ENT != "Nacional", ]
datos_filtrados <- na.omit(datos_filtrados[, c("NOM_ENT", "POB_TOT", variables_estudio)])

poblacion_afectada_municipal <- datos_filtrados[, c("NOM_ENT")]

for (variable in variables_estudio) {
  poblacion_afectada_municipal[[variable]] <-
    datos_filtrados$POB_TOT * (datos_filtrados[[variable]] / 100)
}

entidades_unicas <- unique(poblacion_afectada_municipal$NOM_ENT)

resumen_final <- data.frame(
  Entidad = entidades_unicas,
  stringsAsFactors = FALSE
)

for (estado in entidades_unicas) {
  datos_estado <- subset(poblacion_afectada_municipal, NOM_ENT == estado)

  for (variable in variables_estudio) {
    suma_afectados <- sum(datos_estado[[variable]])
    resumen_final[resumen_final$Entidad == estado, variable] <- suma_afectados
  }
}

resumen_final <- resumen_final[order(resumen_final$Entidad), ]

knitr::kable(resumen_final, digits=0)

```

| | Entidad | ANALF | SBASC | OVSDE | OVSEE | OVSAE | OVPT | VHAC | PL.5000 | PO2SM |
|----|----------------------|--------|---------|--------|-------|--------|--------|---------|---------|----------|
| 1 | Aguascalientes | 30409 | 338193 | 4978 | 3308 | 7881 | 11043 | 190691 | 303272 | 844773 |
| 2 | Baja California | 69638 | 933261 | 7544 | 22043 | 79317 | 72172 | 552569 | 318680 | 2770294 |
| 3 | Baja California Sur | 18703 | 192007 | 3358 | 7682 | 43063 | 40403 | 150163 | 82080 | 366528 |
| 4 | Campeche | 55540 | 279216 | 23285 | 9689 | 36943 | 24929 | 280333 | 277802 | 657231 |
| 7 | Chiapas | 788548 | 2719809 | 136013 | 99988 | 591491 | 688798 | 2061314 | 3195198 | 4816956 |
| 8 | Chihuahua | 103988 | 1036329 | 53137 | 62304 | 62156 | 81853 | 518501 | 538301 | 2514326 |
| 9 | Ciudad de México | 132796 | 1635049 | 4783 | 4769 | 114266 | 58318 | 1367722 | 92835 | 5225821 |
| 5 | Coahuila de Zaragoza | 52880 | 678081 | 9406 | 5347 | 29588 | 23526 | 425446 | 315924 | 1895234 |
| 6 | Colima | 24962 | 204624 | 1967 | 2397 | 4844 | 19178 | 113587 | 98713 | 436999 |
| 10 | Durango | 52734 | 509945 | 52619 | 39022 | 42994 | 78715 | 305560 | 595580 | 1281955 |
| 11 | Guanajuato | 328319 | 2072813 | 118797 | 23169 | 179535 | 149182 | 1051032 | 2051324 | 4185577 |
| 12 | Guerrero | 458108 | 1532264 | 331671 | 49950 | 408471 | 540173 | 1191502 | 1704669 | 2879783 |
| 13 | Hidalgo | 209342 | 931871 | 58513 | 19692 | 119216 | 87793 | 581696 | 1659471 | 2314361 |
| 14 | Jalisco | 245817 | 2485912 | 47873 | 25160 | 63292 | 139091 | 1181627 | 1349320 | 4718807 |
| 15 | México | 507525 | 4292528 | 205825 | 44752 | 473753 | 354215 | 3571690 | 3253663 | 11277406 |
| 16 | Michoacán de Ocampo | 338371 | 2026250 | 65008 | 23283 | 133989 | 265691 | 905739 | 1782054 | 3389682 |
| 17 | Morelos | 88547 | 546420 | 14486 | 5856 | 83183 | 80922 | 384302 | 515167 | 1460087 |
| 18 | Nayarit | 58254 | 364567 | 49970 | 26749 | 37736 | 49947 | 229857 | 450580 | 814272 |
| 19 | Nuevo León | 84962 | 1105988 | 5868 | 6324 | 40588 | 45519 | 762685 | 297175 | 2727860 |

| | Entidad | ANALF | SBASC | OVSDE | OVSEEO | OVSAE | OVPT | VHAC | PL.5000 | PO2SM |
|----|---------------------------------|--------|---------|--------|--------|--------|--------|---------|---------|---------|
| 20 | Oaxaca | 502384 | 1896352 | 80095 | 79514 | 413024 | 578154 | 1240886 | 2454557 | 3333559 |
| 21 | Puebla | 473732 | 2471385 | 74522 | 40387 | 306607 | 351049 | 1724331 | 2375402 | 5158797 |
| 22 | Querétaro | 84704 | 569142 | 39283 | 11108 | 50397 | 47857 | 368013 | 682795 | 1373843 |
| 23 | Quintana Roo | 57868 | 420132 | 21014 | 15279 | 30745 | 44867 | 492229 | 213056 | 1077229 |
| 24 | San Luis Potosí | 144613 | 837509 | 42819 | 33330 | 215529 | 140118 | 480985 | 1016808 | 1904517 |
| 25 | Sinaloa | 108592 | 876906 | 41573 | 12058 | 42435 | 68494 | 574670 | 885671 | 1891665 |
| 26 | Sonora | 58849 | 661148 | 19330 | 19507 | 38178 | 71349 | 494036 | 446306 | 1879281 |
| 27 | Tabasco | 123323 | 705192 | 42265 | 11821 | 148424 | 80901 | 628418 | 1279828 | 1735564 |
| 28 | Tamaulipas | 91593 | 913692 | 8698 | 15100 | 57237 | 49975 | 618302 | 396536 | 2694186 |
| 29 | Tlaxcala | 45364 | 363489 | 11981 | 5488 | 12012 | 23966 | 303910 | 431505 | 1062747 |
| 30 | Veracruz de Ignacio de la Llave | 705271 | 3258176 | 101362 | 86087 | 699118 | 482242 | 1913285 | 3697636 | 6363210 |
| 31 | Yucatán | 143724 | 744201 | 136196 | 15968 | 24673 | 32761 | 628056 | 541165 | 1637759 |
| 32 | Zacatecas | 61534 | 525952 | 37538 | 7464 | 29743 | 20213 | 266974 | 720615 | 1185994 |

e) Comparación Índice de Marginación

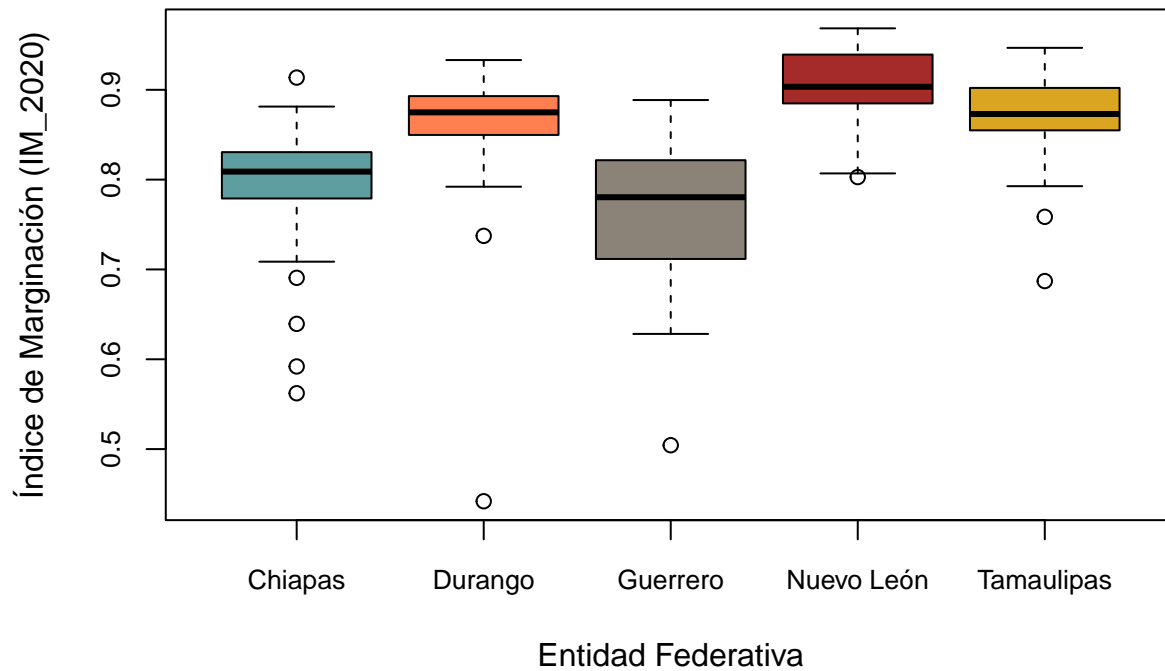
```

entidades_boxplot <- c("Chiapas", "Guerrero", "Nuevo León", "Tamaulipas", "Durango")

datos_boxplot <- subset(marginacion_base, NOM_ENT %in% entidades_boxplot)
datos_boxplot <- na.omit(datos_boxplot[, c("NOM_ENT", variable_imn)])
boxplot(IMN_2020 ~ NOM_ENT,
        data = datos_boxplot,
        main = "Comparación del Índice de Marginación Municipal (IM_2020)",
        xlab = "Entidad Federativa",
        ylab = "Índice de Marginación (IM_2020)",
        col = colores,
        cex.axis = 0.8,
        outline = TRUE)

```

Comparación del Índice de Marginación Municipal (IM_2020)



Observaciones

- Chiapas y Guerrero. La gran mayoría de sus municipios se agrupa en índices altos de marginación. Son los estados con la peor situación generalizada a nivel municipal.
- Nuevo León. La mediana de marginación se encuentra en los valores más bajos. Posibles atípicos hacia la alta marginación, serían municipios rurales aislados. Baja Marginación Concentrada. Sus municipios son, en su mayoría, los menos marginados.
- Tamaulipas y Durango. Posición intermedia, se observan algunos valores atípicos de alta marginación. Coexisten municipios de baja marginación ciudades principales y municipios con problemas severos zonas rurales o fronterizas.

Distribuciones de Probabilidad

1. Ejercicio concurso de TV

El 30% de un determinado pueblo ve un concurso que hay en televisión. Desde el concurso se llama por teléfono a 10 personas del pueblo elegidas al azar. Determinar la probabilidad que, de las 10 personas elegidas:

Datos:

- Número de eventos $n = 10$
- Probabilidad de éxito $p = 0.30$

El ejercicio corresponde a una Distribución Binomial, ya que cumple con las siguientes condiciones:

- Hay un número fijo de ensayos $n = 10$ personas.
- Cada ensayo tiene solo dos resultados posibles: ver el programa (éxito) o no verlo (fracaso).
- La probabilidad de éxito es constante en cada ensayo $p = 0.30$
- Los ensayos son independientes.

a) Más de ocho personas estén viendo el programa

$$P(X = 9) + P(X = 10)$$

```
prob_mas_de_ocho <- dbinom(x = 9, size = 10, prob = 0.30) +  
  dbinom(x = 10, size = 10, prob = 0.30)  
prob_mas_de_ocho  
  
## [1] 0.0001436859
```

$$P(X > 8) = 0.0001436859$$

La probabilidad de que, al llamar a 10 personas al azar, más de 8, es decir, 9 o 10 estén viendo el programa es de aproximadamente 0.015%

b) Al menos una persona de las diez esté viendo el programa

$$P(X \geq 1) = 1 - P(X < 1)$$

```
prob_al_menos_1 <- 1 - pbinom(q = 0, size = 10, prob = 0.30)  
prob_al_menos_1  
  
## [1] 0.9717525
```

$$P(X \geq 1) = 0.9717525$$

La probabilidad de que al menos una persona de las diez elegidas al azar esté viendo el programa es de aproximadamente 97.18%.

2. Ejercicio células de levadura

Se estudia la distribución de células de levadura en 400 cuadrículas de un hemacitómetro. Este es un aparato semejante al utilizado para hacer los recuentos de las células sanguíneas y de otras estructuras microscópicas suspendidas en líquidos. ¿Cuál es la probabilidad de encontrar dos células en las cuadrículas observadas? Después de haber muestreado 400 cuadrículas y contado la cantidad de células de levadura en cada una de ellas, se obtuvo la siguiente tabla de frecuencias:

| Cantidad de celulas por cuadrícula | Cantidad de cuadrículas observadas |
|------------------------------------|------------------------------------|
| 0 | 75 |
| 1 | 103 |
| 2 | 121 |
| 3 | 54 |
| 4 | 30 |
| 5 | 13 |
| 6 | 2 |
| 7 | 1 |
| 8 | 1 |

a) Calcular el valor esperado de células por cuadrícula $E(X)$

```
conteo_celulas <- c(0, 1, 2, 3, 4, 5, 6, 7, 8)
frecuencia <- c(75, 103, 121, 54, 30, 13, 2, 1, 1)
sum_celulas_por_frecuencia <- sum(conteo_celulas * frecuencia)
total_cuadriculas <- sum(frecuencia)
```

```
E_X <- sum_celulas_por_frecuencia / total_cuadriculas
E_X
```

```
## [1] 1.7975
```

En promedio hay 1.79 células por cuadrícula.

b) ¿Cuál es la probabilidad de encontrar dos células en las cuadrículas observadas?

```
lambda <- E_X
x_valor <- 2

prob_dos_celulas <- dpois(x = x_valor, lambda = lambda)
prob_dos_celulas
```

```
## [1] 0.2677093
```

Utilizando el modelo de Poisson la probabilidad de encontrar exactamente dos células en una cuadrícula elegida al azar es del 26.77%.

c) Calcular la probabilidad de que los conteos de células se encuentren entre 3 y 7.

Esto significa que queremos la probabilidad de que X sea mayor o igual a 3 y menor o igual a 7

$$P(3 \leq X \leq 7)$$

$$P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)$$

```
lambda <- E_X

prob_3 <- dpois(x = 3, lambda = lambda)
prob_4 <- dpois(x = 4, lambda = lambda)
prob_5 <- dpois(x = 5, lambda = lambda)
prob_6 <- dpois(x = 6, lambda = lambda)
prob_7 <- dpois(x = 7, lambda = lambda)

prob_rango_3_a_7 <- prob_3 + prob_4 + prob_5 + prob_6 + prob_7
prob_rango_3_a_7
```

```
## [1] 0.268153
```

La probabilidad de que el conteo esté entre 3 y 7 células es de 26.81%.

3. Ejercicio tornillos

La longitud X de ciertos tornillos es una variable aleatoria con distribución normal de media 30mm y desviación típica 0.2mm. Se aceptan como validos aquellos que cumplen $29.5\text{mm} < X < 30.4\text{mm}$. Calcular las siguientes probabilidades:

Datos:

- Media (μ): 30 mm
- Desviación Típica (σ): 0.2 mm
- Rango Aceptable: $29.5\text{ mm} < X < 30.4\text{ mm}$

a) Tornillos no aceptables por cortos

$$P(X \leq 29.5)$$

```
media <- 30
desviacion_tipica <- 0.2
limite_corto <- 29.5
prob_cortos <- pnorm(q = limite_corto, mean = media, sd = desviacion_tipica)
prob_cortos

## [1] 0.006209665
```

$$P(X \leq 29.5) : 0.0062$$

Probabilidad de tornillos no aceptables por cortos: 0.62%

b) Tornillos no aceptables por largos Los tornillos no son aceptables por largos si su longitud

$$P(X \geq 30.4) = 1 - P(X < 30.4)$$

```
limite_largo <- 30.4
prob_largos <- 1 - pnorm(q = limite_largo, mean = media, sd = desviacion_tipica)
prob_largos
```

```
## [1] 0.02275013
```

La probabilidad de que un tornillo sea rechazado por ser demasiado largo, más de 30.4 mm, es del 2.275%.

c) Tornillos no validos

Los tornillos no válidos son aquellos que son cortos O largos:

$$P(\text{NoValidos}) = P(X \leq 29.5) + P(X \geq 30.4)$$

```
prob_no_validos <- prob_cortos + prob_largos
prob_validos <- pnorm(limite_largo, media, desviacion_tipica) -
  pnorm(limite_corto, media, desviacion_tipica)

prob_no_validos_alternativa <- 1 - prob_validos
prob_no_validos_alternativa
```

```
## [1] 0.0289598
```

Probabilidad de tornillos no validos: 0.0289598 o 2.895%

d) ¿Cual es la probabilidad que de una muestra de 15 tornillos entre 13 y 15 tornillos resulten validos?

$P(13 \leq Y \leq 15)$, donde Y es el número de tornillos válidos.

$$P(Y = 13) + P(Y = 14) + P(Y = 15)$$

```
n_muestra <- 15
p_valido <- prob_validos

prob_13 <- dbinom(x = 13, size = n_muestra, prob = p_valido)
prob_14 <- dbinom(x = 14, size = n_muestra, prob = p_valido)
```

```
prob_15 <- dbinom(x = 15, size = n_muestra, prob = p_valido)
```

```
prob_rango_13_a_15 <- prob_13 + prob_14 + prob_15
```

```
print(paste("Probabilidad (Y=13):", prob_13))
```

```
## [1] "Probabilidad (Y=13): 0.0600985611396816"
```

```
print(paste("Probabilidad (Y=14):", prob_14))
```

```
## [1] "Probabilidad (Y=14): 0.28787750352453"
```

```
print(paste("Probabilidad (Y=15):", prob_15, 4))
```

```
## [1] "Probabilidad (Y=15): 0.643514240882078 4"
```

```
prob_rango_13_a_15
```

```
## [1] 0.9914903
```

Probabilidad ($13 \leq Y \leq 15$): 0.99149 o 99.149%

4. Ejercicio Computadoras

Se tiene conocido por estudios previos que el tiempo de vida útil de computadoras personales de cierta marca sigue una distribución $N(1000, 20)$ días.

Datos (Distribución Normal):

- Tiempo de vida útil (X): Sigue una $N(\mu, \sigma^2)$
- Media (μ): 1000 días
- Desviación Típica (σ): 20 días

a) ¿A partir de cuantas horas se espera que fallen minimo el 95% de las computadoras?

Encontrar el valor x tal que $P(X \leq x) = 0.95$

```
media <- 1000
```

```
desviacion_tipica <- 20
```

```
probabilidad <- 0.95
```

```
dias_95_percentil <- qnorm(p = probabilidad, mean = media, sd = desviacion_tipica)
```

```
dias_95_percentil
```

```
## [1] 1032.897
```

```
horas_95_percentil <- dias_95_percentil * 24
```

```
horas_95_percentil
```

```
## [1] 24789.53
```

Percentil 95 tiempo a partir del cual han fallado al menos el 95%:

- 1032.897 días
- 24789.53 horas

b) La garantía determina que se reemplazara una computadora si dura sin fallar menos de 800 días. Determinar cual es la probabilidad de que reemplacen los equipos a cierto cliente.

Una computadora se reemplaza si falla en menos de 800 días $P(X < 800)$

```

limite_garantia <- 800
prob_reemplazo <- pnorm(q = limite_garantia, mean = media, sd = desviacion_tipica)
prob_reemplazo

```

```
## [1] 7.619853e-24
```

Probabilidad de reemplazo ($P(X < 800 \text{ días})$): 7.619853e-24 es decir cercano a cero.

c) Dado un lote de 20 computadoras, ¿cual es numero esperado de computadoras que se regresen?

```

n_lote_20 <- 20
p_regreso <- prob_reemplazo
valor_esperado_regresos <- n_lote_20 * p_regreso
valor_esperado_regresos

```

```
## [1] 1.523971e-22
```

Número Esperado de regresos ($E(Y)$): 1.523971e-22 practicamente cero.

d) ¿Cual es la probabilidad de que se regresen entre una y dos computadoras de un lote de 30 computadoras?

$$P(1 \leq Y \leq 2) = P(Y = 1) + P(Y = 2)$$

```

n_lote_30 <- 30
p_regreso <- prob_reemplazo

prob_1 <- dbinom(x = 1, size = n_lote_30, prob = p_regreso)
prob_2 <- dbinom(x = 2, size = n_lote_30, prob = p_regreso)

prob_1_o_2_regresadas <- prob_1 + prob_2
prob_1_o_2_regresadas

```

```
## [1] 2.285956e-22
```

Probabilidad de 1 o 2 regresos: 2.285956e-22 también cercano a cero.