**Actividad final correspondiente al curso de Estadística y Probabilidad  
(Parte 1)**

Luis Neumann

5 de Junio del 2020.

**Resumen**

El presente reporte, tiene como objetivo mostrar los resultados de la estadística descriptiva aplicada a una muestra de 500 combos de hamburguesa, salsa, papas y refresco, obtenida de la siguiente URL:

[*https://codap.concord.org/releases/latest/static/dg/en/cert/index.html#shared=106296*](https://codap.concord.org/releases/latest/static/dg/en/cert/index.html#shared=106296)

**Palabras clave:** estadística, descripción, 500, combos, hamburguesa, carne, salsa, papas, refresco

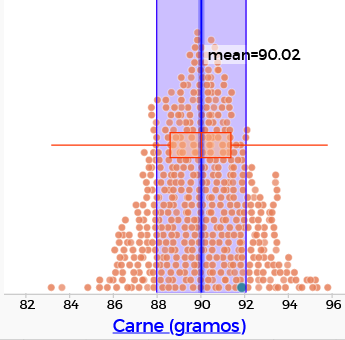
# Descripción del comportamiento de la variables cuantitativas en términos de tendencia, variabilidad y distribución

## Variable “Carne”1

**Tabla 1.** Resumen de estadística descriptiva de la variable “Carne”

|  |  |
| --- | --- |
| ***Carne (grs.)*** | |
|  |  |
| Media | 90.021807 |
| Error típico | 0.091332354 |
| Mediana | 90.02615 |
| Moda | 90.0396 |
| Desviación estándar | 2.042253524 |
| Varianza de la muestra | 4.170799456 |
| Curtosis | -0.077008602 |
| Coeficiente de asimetría | 0.008846088 |
| Rango | 12.637 |
| Mínimo | 83.1616 |
| Máximo | 95.7986 |
| Suma | 45010.9035 |
| Cuenta | 500 |
| Mayor (1) | 95.7986 |
| Menor (1) | 83.1616 |
| Nivel de confianza (95.0%) | 0.179443361 |

**Figura 1.** Distibución de la variable “Carne”.



Como es posible observar, la variable “**Carne**” presenta una distribución aproximada a la distibución normal. Los valores giran alrededor de la media aritmética correspondiente a los **90.02** (grs.). Los valores tienen una desviación estándar aproximada de **2.04** (grs.) y un coeficiente de variación de aproximadamente el **2**%.

El primer cuartil se presenta a los **88.59** (grs.) y el tercer cuartil a los **91.37** (grs.) No se observan datos atípicos en la gráfica.

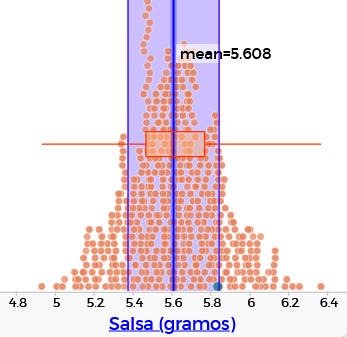
Debido a que el coeficiente de variación es pequeño, podemos encontrar que hay poca variabilidad en los datos y el centro (la media aritmética) los representa muy bien a todos ellos.

## Variable “Salsa”1

**Tabla 2.** Resumen de estadística descriptiva de la variable “Salsa”

|  |  |
| --- | --- |
| ***Salsa (grs.)*** | |
|  |  |
| Media | 5.60811004 |
| Error típico | 0.010476723 |
| Mediana | 5.60025 |
| Moda | 5.48999 |
| Desviación estándar | 0.234266648 |
| Varianza de la muestra | 0.054880862 |
| Curtosis | -0.015408673 |
| Coeficiente de asimetría | 0.099845103 |
| Rango | 1.43587 |
| Mínimo | 4.92999 |
| Máximo | 6.36586 |
| Suma | 2804.05502 |
| Cuenta | 500 |
| Mayor (1) | 6.36586 |
| Menor (1) | 4.92999 |
| Nivel de confianza (95.0%) | 0.020583926 |

**Figura 2.** Distibución de la variable “Salsa”.



Respecto a la variable “**Salsa**”, es posible observar que, al igual que la variable “Carne”, presenta uns distribución parecida a la normal. La mayor parte de los datos se concentra cerca de la media aritmética de **5.608** (grs) y tiene pocos datos muy por debajo de la media y pocos datos muy por encima de la media. La desviación estándar de los datos es de aproximadamente **0.23** (grs.) y presentan un coeficiente de variación de poco más del **4**(%). Se observa que el primer cuantil se presenta a los **5.46** (grs.) y el tercer cuantil a los **5.76** (grs.). No se observan datos atípicos en la gráfica.

Las medidas estadísticas presentadas, nos indican que se trata de un conjunto de valores que presentan muy poca variabilidad y la media aritmética es un valor que los representa de manera satisfactoria a todos ellos.

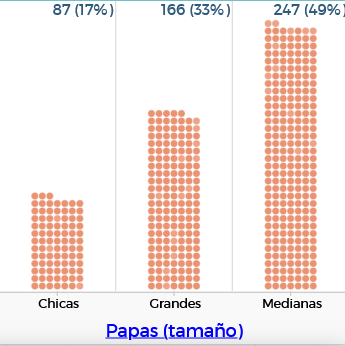
# Descripción del comportamiento de la variables cualitativas

## Variable “Papas”1

**Tabla 3.** Resumen de estadística descriptiva de la variable “Papas”

|  |  |  |
| --- | --- | --- |
| **Etiquetas de fila** | **Cuenta de Papas** |  |
| Chicas | 87 |  |
| Grandes | 166 |  |
| Medianas | 247 |  |
| **Total general** | **500** |  |
|  |  |  |
|  |  |  |
| **Papas** | **Frecuencia** | **%** |
| **Chicas** | 87 | 17% |
| **Medianas** | 247 | 49% |
| **Grandes** | 166 | 33% |
| **Total** | 500 | 100% |

**Figura 3.** Distibución de la variable “Papas”.



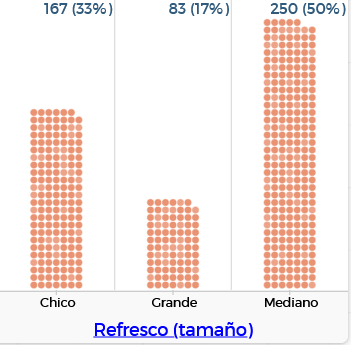
Se encontró que la variable “Papas” se puede dividir en tres categorías en función de su tamaño del recipiente que las contiene y la cantidad de producto que contiene: “Chicas”, “Grandes” y “Medianas”, siendo estas últimas las más pedidas por el consumidor. La presentación menos solicitada es la presentación correspondiente a las “Chicas”.

## Variable “Refresco”1

**Tabla 4.** Resumen de estadística descriptiva de la variable “Refresco”

|  |  |  |
| --- | --- | --- |
| **Etiquetas de fila** | **Cuenta de Refresco** |  |
| Chico | 167 |  |
| Grande | 83 |  |
| Mediano | 250 |  |
| **Total general** | **500** |  |
|  |  |  |
|  |  |  |
| **Refresco** | **Frecuencia** | **%** |
| **Chico** | 167 | 33% |
| **Mediano** | 250 | 50% |
| **Grande** | 83 | 17% |
| **Total** | 500 | 100% |

**Figura 4.** Distibución de la variable “Refresco”.



La variable “Refresco” también presenta tres categorías en función del tamaño del envase y cantidad de producto: “Chico”, “Grande” y “Mediano”, siendo este último el que es solicitado por **50** (%) de los consumidores, le sigue el tamaño “Chico” y finalmente el menos consumido es el tamaño “Grande”.

# Descripción del comportamiento conjunto de las variables “Carne” y “Salsa”

Respecto a la comparación de ambas variables encontramos que se trata del mismo tipo de métrica en ambos casos: la masa. También es posible observar que la distribución de los datos en ambas variables es similar, pocos datos muy por debajo de la media y pocos ejemplares muy por encima de la media aritmética. La mayor parte de datos en cada una de las variables se concentra en torno a la media aritmética. Para el caso de la variable “Carne”, el coeficiente de variación es de 0.023 y para la variable “Salsa” es de 0.042, por lo que podemos decir que para la métrica “masa”, la variable “Salsa” presenta mayor variabilidad que la variable “Carne”.

**Figura 5.** Diagrama de dispersión Carne Vs Salsa

|  |  |
| --- | --- |
| **Coeficiente de correlación:** | 0.88167831 |

El diagrama de dispersión de la Figura 5 muestra que los puntos se concentran en una región del plano cartesiano, de tal forma que pudieran estar distribuidos a lo largo de una recta con pendiente positiva y ordenada al origen mayor a cero.

Al calcular la ecuación de dicha recta por el método de recta de ajuste por mínimos cuadrados, obtenemos:

Y con un coeficiente de correlación cercano a 1, lo que nos indica que los datos se ajustan de buena manera a una recta de pendiente positiva y que ambas variables “varían” más o menos de la misma forma.

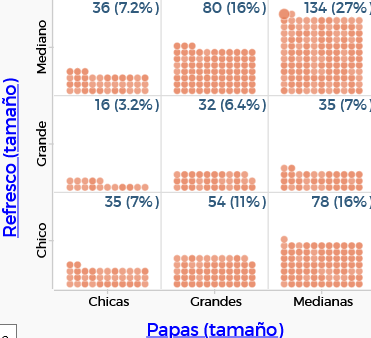
Este modelo matemático nos permite realizar buenas estimaciones de la masa de la salsa, en función de la masa de la carne para valores cercanos a los 83(grs.) y 96(grs.).

# Descripción del comportamiento conjunto de las variables “Papas” y “Refresco”

**Tabla 5.** Tabla de doble entrada, tabla cruzada o tabla de contingencia de las variables cualitativas “Papas” y “Refresco”

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Conteo* | **Refresco** | | |  |
| **Papas** | **Chico** | **Mediano** | **Grande** | **Total** |
| **Chicas** | 35 | 36 | 16 | **87** |
| **Medianas** | 78 | 134 | 35 | **247** |
| **Grandes** | 54 | 80 | 32 | **166** |
| Total | **167** | **250** | **83** | 500 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| *Frecuencias relativas* | **Refresco** | | |  |
| **Papas** | **Chico** | **Mediano** | **Grande** | **Total** |
| **Chicas** | 0.07 | 0.072 | 0.032 | **0.174** |
| **Medianas** | 0.156 | 0.268 | 0.07 | **0.494** |
| **Grandes** | 0.108 | 0.16 | 0.064 | **0.332** |
| Total | **0.334** | **0.5** | **0.166** | 1 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| *Porcentajes* | **Refresco** | | |  |
| **Papas** | **Chico** | **Mediano** | **Grande** | **Total** |
| **Chicas** | 7% | 7% | 3% | **17%** |
| **Medianas** | 16% | 27% | 7% | **49%** |
| **Grandes** | 11% | 16% | 6% | **33%** |
| Total | **33%** | **50%** | **17%** | 100% |

**Figura 6.** Diagrama Papas Vs Refresco



Las variables cualitativas “Papas” y “Refresco” se presentan de manera simultánea en cada una de las observaciones. La representación de la tabla de contingencia, nos permite analizar las variables tanto por separado como de manera conjunta. Al final de cada tabla, se aprecia el total de observaciones para cada tamaño de papas y para cada tamaño de refresco, mientras que en los cruces, se tiene los totales para las dos variables tomadas de manera conjunta en cada una de las posibles combinaciones de tamaño de papas y refresco.

De las combinanciones mencionadas anteriormente, es posible destacar que la más común es la que corresponde a las papas medianas y refresco mediano con un 26.8 (%) y la menos común es papas chicas con refresco grande con un **3.2** (%)