

Informe de Análisis Exploratorio de **mtcars**

Melani Forsythe Matos
Daniela Guerrero Álvarez
Rubén Martínez Rojas

El conjunto de datos `mtcars` en R es un conjunto clásico que contiene datos sobre automóviles, y tiene 32 observaciones (filas) de 11 variables (columnas). A continuación, describo cada una de las variables, su significado, el tipo de escala, y si son discretas o continuas:

1. mpg (Miles per Gallon)

- **Descripción:** Consumo de combustible del automóvil en millas por galón.
- **Escala:** Cuantitativa Continua.
- **Significado:** Representa la eficiencia del combustible del automóvil, es decir, cuántas millas puede recorrer el automóvil por cada galón de gasolina.

Media	Mediana	Moda	Simetría	Curtosis
20.090625	19.200000	10.400000	0.610655	2.627234

Figure 1: Medidas de tendencia central para la variable `mpg`.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
36.324103	6.026948	23.500000	29.998808

Figure 2: Medidas de dispersión para la variable `mpg`.

- **Media:** El promedio de millas por galón es de 20.09. Este valor nos indica el rendimiento promedio de combustible de los vehículos analizados.
- **Mediana:** El valor central es de 19.2 mpg, lo que implica que el 50% de los vehículos tienen un rendimiento inferior o igual a este valor.
- **Moda:** El valor que más se repite es 10.4 mpg, lo cual sugiere que un subconjunto de vehículos tiene un rendimiento significativamente bajo.
- **Simetría:** Con un valor de 0.61, la distribución está sesgada positivamente, indicando que hay más valores bajos de mpg con algunos valores altos.

- **Curtosis:** El valor de 2.63 indica que la distribución es leptocúrtica, con colas más pesadas que una distribución normal.
- **Varianza y Desviación Estándar:** La varianza es 36.32 y la desviación estándar es 6.03, lo que revela una dispersión moderada alrededor de la media.
- **Rango:** La diferencia entre el máximo y mínimo es de 23.5 mpg, mostrando una considerable variabilidad en el rendimiento.
- **Coefficiente de Variación:** Con un 30%, hay una variabilidad moderada en relación con la media.

2. cyl (Cylinders)

- **Descripción:** Número de cilindros en el motor del automóvil.
- **Escala:** Cuantitativa Discreta.
- **Significado:** Indica cuántos cilindros tiene el motor. Generalmente, los valores comunes son 4, 6 u 8 cilindros.

Media	Mediana	Moda	Simetría	Curtosis
6.1875000	6.0000000	8.0000000	-0.1746119	1.2378802

Figure 3: Medidas de tendencia central para la variable cyl.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
3.189516	1.785922	4.000000	28.863380

Figure 4: Medidas de dispersión para la variable cyl.

- **Media:** El número promedio de cilindros es 6.19, indicando que la mayoría de los vehículos tienen entre 6 y 8 cilindros.
- **Mediana:** El valor central es 6, lo que significa que la mitad de los autos tiene 6 cilindros o menos.
- **Moda:** La moda es 8 cilindros, lo que sugiere que es el tipo más común en el conjunto de datos.
- **Simetría:** Un valor de -0.17 indica una ligera asimetría negativa, donde algunos vehículos tienen menos cilindros de lo esperado.

- **Curtosis:** Con 1.24, la distribución es platicúrtica, más plana que una normal.
- **Varianza y Desviación Estándar:** La varianza es 3.19 y la desviación estándar es 1.79, sugiriendo una dispersión moderada.
- **Rango:** El rango de 4 cilindros refleja una variabilidad moderada en la configuración del motor.
- **Coefficiente de Variación:** Con 28.86%, indica una moderada variabilidad relativa en el número de cilindros.

3. disp (Displacement)

- **Descripción:** Desplazamiento del motor en pulgadas cúbicas.
- **Escala:** Cuantitativa Continua.
- **Significado:** Es el volumen total desplazado por todos los pistones dentro del motor en una sola revolución. Es una medida del tamaño del motor.

Media	Mediana	Moda	Simetría	Curtosis
230.721875	196.300000	275.800000	0.381657	1.792788

Figure 5: Medidas de tendencia central para la variable `disp`.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
15360.79983	123.93869	400.90000	53.71779

Figure 6: Medidas de dispersión para la variable `disp`.

- **Media:** El promedio de desplazamiento es 230.72 pulgadas cúbicas, lo que refleja un tamaño del motor significativo en los vehículos.
- **Mediana:** El valor central es 196.3, indicando que la mitad de los vehículos tiene un desplazamiento inferior a este.
- **Moda:** La moda de 275.8 sugiere que algunos vehículos tienen un motor considerablemente grande.
- **Simetría:** Un valor de 0.38 indica una ligera asimetría positiva, con algunos motores muy grandes que aumentan la media.

- **Curtosis:** Con 1.79, la distribución es platicúrtica, indicando colas menos pesadas.
- **Varianza y Desviación Estándar:** La alta varianza de 15360.80 y desviación estándar de 123.94 indican una gran dispersión.
- **Rango:** El rango de 400.9 muestra una gran diferencia en el tamaño del motor entre los vehículos.
- **Coefficiente de Variación:** Con un 53.72%, existe una alta variabilidad relativa en el desplazamiento.

4. hp (Horsepower)

- **Descripción:** Potencia del motor en caballos de fuerza.
- **Escala:** Es una variable Cuantitativa que puede tener valores Continuos, pero en este caso solo guarda valores Discretos, por tanto la trataremos como Cuantitativa Discreta.
- **Significado:** Mide la potencia del motor, es decir, la capacidad del motor para realizar trabajo en una unidad de tiempo.

Media	Mediana	Moda	Simetría	Curtosis
146.6875000	123.0000000	110.0000000	0.7260237	2.8644489

Figure 7: Medidas de tendencia central para la variable **hp**.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
4700.86694	68.56287	283.00000	46.74077

Figure 8: Medidas de dispersión para la variable **hp**.

- **Media:** El promedio es 146.69 caballos de fuerza, reflejando una potencia moderada en los vehículos.
- **Mediana:** Con 123 hp, la mitad de los vehículos tiene menos potencia que este valor.
- **Moda:** La moda es 110 hp, común en vehículos de potencia media.
- **Simetría:** Un valor de 0.73 indica una asimetría positiva, sugiriendo que algunos autos tienen potencia mucho mayor que la media.

- **Curtosis:** El valor de 2.86 indica colas más pesadas, sugiriendo algunos valores extremos de potencia.
- **Varianza y Desviación Estándar:** La varianza es 4700.87 y la desviación estándar 68.56, lo que indica una alta dispersión en la potencia.
- **Rango:** El rango de 283 hp muestra una diferencia significativa entre los autos menos y más potentes.
- **Coefficiente de Variación:** Un 46.74% indica una alta variabilidad relativa.

5. drat (Rear Axle Ratio)

- **Descripción:** Relación de transmisión del eje trasero.
- **Escala:** Cuantitativa Continua.
- **Significado:** Es la relación entre las revoluciones del eje de transmisión y las revoluciones del eje trasero. Afecta el rendimiento y la velocidad del vehículo.

Media	Mediana	Moda	Simetría	Curtosis
3.5965625	3.6950000	3.0700000	0.2659039	2.2852994

Figure 9: Medidas de tendencia central para la variable **drat**.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
0.2858814	0.5346787	2.1700000	14.8663824

Figure 10: Medidas de dispersión para la variable **drat**.

- **Media:** La relación promedio es 3.60, reflejando las configuraciones de transmisión en los vehículos.
- **Mediana:** Con 3.70, la mitad de las relaciones del eje están por debajo de este valor.
- **Moda:** El valor más común es 3.07.
- **Simetría:** Un valor de 0.27 sugiere una ligera asimetría positiva.

- **Curtosis:** Con 2.29, la distribución es mesocúrtica, cercana a la normal.
- **Varianza y Desviación Estándar:** La baja varianza de 0.29 y desviación estándar de 0.53 indican poca dispersión.
- **Rango:** Un rango de 2.17 muestra cierta variabilidad en las relaciones del eje.
- **Coefficiente de Variación:** El 14.87% sugiere baja variabilidad relativa.

6. wt (Weight)

- **Descripción:** Peso del automóvil en miles de libras.
- **Escala:** Cuantitativa Continua.
- **Significado:** El peso del automóvil influye en su eficiencia, aceleración y manejo.

Media	Mediana	Moda	Simetría	Curtosis
3.2172500	3.3250000	3.4400000	0.4231465	2.9772892

Figure 11: Medidas de tendencia central para la variable **wt**.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
0.9573790	0.9784574	3.9110000	30.4128508

Figure 12: Medidas de dispersión para la variable **wt**.

- **Media:** El peso promedio es 3.22 (en miles de libras), reflejando el peso general de los vehículos.
- **Mediana:** El valor central es 3.33.
- **Moda:** La moda es 3.44, indicando que es un peso común entre los vehículos.
- **Simetría:** Un valor de 0.42 sugiere una ligera asimetría positiva.
- **Curtosis:** Con 2.98, la distribución es leptocúrtica.
- **Varianza y Desviación Estándar:** La varianza de 0.96 y desviación estándar de 0.98 indican una moderada dispersión.

- **Rango:** Un rango de 3.91 refleja una diferencia considerable en el peso de los vehículos.
- **Coefficiente de Variación:** Con 30.41%, hay una variabilidad moderada en relación a la media.

7. qsec (1/4 Mile Time)

- **Descripción:** Tiempo en segundos para recorrer un cuarto de milla.
- **Escala:** Cuantitativa Continua.
- **Significado:** Es una medida del tiempo que tarda el automóvil en recorrer un cuarto de milla, comúnmente usado para medir el rendimiento en aceleración.

Media	Mediana	Moda	Simetría	Curtosis
17.8487500	17.7100000	17.0200000	0.3690453	3.3351142

Figure 13: Medidas de tendencia central para la variable qsec.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
3.193166	1.786943	8.400000	10.011588

Figure 14: Medidas de dispersión para la variable qsec.

- **Media:** El promedio es 17.85 segundos, indicando el tiempo medio que tardan los vehículos en recorrer un cuarto de milla.
- **Mediana:** El valor central es 17.71 segundos.
- **Moda:** La moda es 17.02 segundos.
- **Simetría:** Un valor de 0.37 indica una ligera asimetría positiva.
- **Curtosis:** Con 3.34, la distribución es leptocúrtica.
- **Varianza y Desviación Estándar:** La varianza es 3.19 y la desviación estándar es 1.79, lo que indica una dispersión moderada.
- **Rango:** El rango de 8.4 segundos muestra variabilidad en la aceleración de los vehículos.
- **Coefficiente de Variación:** Con un 10.01%, hay poca variabilidad relativa.

8. vs (Engine Shape)

- **Descripción:** Forma del motor (0 = motor en V, 1 = motor en línea).
- **Escala:** Esta es de cierta manera una variable cualitativa nominal, lo que esta convertida a Cuantitativa Discreta (Binaria).
- **Significado:** Indica el tipo de configuración del motor: si los cilindros están dispuestos en forma de V o en línea.

Media	Mediana	Moda	Simetría	Curtosis
0.4375000	0.0000000	0.0000000	0.2402577	0.9980624

Figure 15: Medidas de tendencia central para la variable vs.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
0.2540323	0.5040161	1.0000000	115.2036866

Figure 16: Medidas de dispersión para la variable vs.

- **Media:** El promedio es 0.44, lo que sugiere que aproximadamente el 44% de los vehículos tienen un motor en línea.
- **Mediana:** La mediana es 0, lo que significa que la mayoría de los vehículos tienen motores en V.
- **Moda:** La moda es 0, reforzando que la configuración de motor en V es la más común.
- **Simetría:** Un valor de 0.24 indica una ligera asimetría positiva.
- **Curtosis:** Con 0.99, la distribución es casi mesocúrtica.
- **Varianza y Desviación Estándar:** La varianza es 0.25 y la desviación estándar es 0.50, indicando poca dispersión.
- **Rango:** El rango es 1, mostrando la variabilidad binaria.
- **Coefficiente de Variación:** Con 115.20%, existe una alta variabilidad relativa en esta variable binaria.

9. am (Transmission)

- **Descripción:** Tipo de transmisión (0 = automática, 1 = manual).
- **Escala:** Esta es de cierta manera una variable cualitativa nominal, lo que esta convertida a CUantitativa Discreta (Binaria).
- **Significado:** Indica si el automóvil tiene una transmisión automática o manual.

Media	Mediana	Moda	Simetría	Curtosis
0.4062500	0.0000000	0.0000000	0.3640159	1.0752586

Figure 17: Medidas de tendencia central para la variable **am**.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
0.2489919	0.4989909	1.0000000	122.8285335

Figure 18: Medidas de dispersión para la variable **am**.

- **Media:** El promedio es 0.41, indicando que aproximadamente el 41% de los vehículos tienen transmisión manual.
- **Mediana:** El valor central es 0, indicando que la mayoría de los vehículos tienen transmisión automática.
- **Moda:** La moda es 0, reforzando que la configuración automática es la más común.
- **Simetría:** Un valor de 0.36 indica una ligera asimetría positiva.
- **Curtosis:** Con 1.08, la distribución es bastante plana.
- **Varianza y Desviación Estándar:** La varianza de 0.25 y desviación estándar de 0.50 indican poca dispersión.
- **Rango:** Un rango de 1 refleja la variabilidad binaria.
- **Coefficiente de Variación:** Con 122.83%, hay una alta variabilidad relativa.

10. gear (Gears)

- **Descripción:** Número de velocidades de la caja de cambios.
- **Escala:** Cuantitativa Discreta.

- **Significado:** Indica cuántas marchas tiene la caja de cambios del automóvil.

Media	Mediana	Moda	Simetría	Curtosis
3.6875000	4.0000000	3.0000000	0.5288545	1.9302493

Figure 19: Medidas de tendencia central para la variable `gear`.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
0.5443548	0.7378041	2.0000000	20.0082458

Figure 20: Medidas de dispersión para la variable `gear`.

- **Media:** El promedio es 3.69, indicando un predominio de autos con 3 a 4 marchas.
- **Mediana:** El valor central es 4 marchas.
- **Moda:** La moda es 3 marchas, lo que sugiere que es la configuración más común.
- **Simetría:** Un valor de 0.53 indica una asimetría positiva, con algunos vehículos que tienen más marchas.
- **Curtosis:** Con 1.93, la distribución es mesocúrtica.
- **Varianza y Desviación Estándar:** La varianza de 0.54 y desviación estándar de 0.74 sugieren dispersión moderada.
- **Rango:** El rango de 2 marchas muestra cierta variabilidad en el número de marchas.
- **Coefficiente de Variación:** Con un 20%, la variabilidad relativa es moderada.

11. `carb` (Carburetors)

- **Descripción:** Número de carburadores.
- **Escala:** Cuantitativa Discreta.
- **Significado:** Indica cuántos carburadores tiene el automóvil, lo que afecta la mezcla de aire y combustible y, por ende, el rendimiento del motor.

Media	Mediana	Moda	Simetría	Curtosis
2.812500	2.000000	2.000000	1.050874	4.257043

Figure 21: Medidas de tendencia central para la variable `carb`.

Varianza	Desviación Estándar	Rango	Coefficiente de Variación
2.608871	1.615200	7.000000	57.429333

Figure 22: Medidas de dispersión para la variable `carb`.

- **Media:** El promedio es 2.81 carburadores, reflejando una moderada complejidad en los motores.
- **Mediana:** El valor central es 2 carburadores.
- **Moda:** La moda es 2 carburadores, indicando que es la configuración más común.
- **Simetría:** Un valor de 1.05 indica una fuerte asimetría positiva, con algunos motores que tienen muchos carburadores.
- **Curtosis:** Con 4.26, la distribución es leptocúrtica, con colas más pesadas.
- **Varianza y Desviación Estándar:** La varianza de 2.61 y desviación estándar de 1.62 indican una dispersión significativa.
- **Rango:** El rango de 7 carburadores muestra una gran diferencia en la complejidad de los motores.
- **Coefficiente de Variación:** Con 57.43%, existe una alta variabilidad relativa.

En general, el análisis muestra que algunas variables, como *mpg*, *hp*, y *disp*, presentan una alta dispersión y asimetría, lo que sugiere que los vehículos en el conjunto de datos tienen una gran diversidad en términos de eficiencia, potencia y tamaño del motor. Otras variables, como *vs* y *am*, son binarias, con poca variabilidad interna. El coeficiente de variación y la simetría son claves para entender cómo los datos están distribuidos y cómo pueden variar las características del vehículo.

1 Resultados de la Matriz de Correlación

A continuación se presentan algunos de los hallazgos clave de la matriz de correlación:

Variable	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
mpg	1.00	-0.85	-0.87	-0.77	0.68	-0.87	0.74	0.66	0.60	0.48	-0.55
cyl	-0.85	1.00	0.90	0.78	-0.50	0.87	-0.59	-0.81	-0.52	0.36	0.55
disp	-0.87	0.90	1.00	0.91	-0.49	0.88	-0.71	-0.71	-0.49	0.38	0.66
hp	-0.77	0.78	0.91	1.00	-0.25	0.66	-0.66	-0.55	-0.68	0.43	0.66
drat	0.68	-0.50	-0.49	-0.25	1.00	-0.39	0.11	0.13	0.79	0.29	-0.09
wt	-0.87	0.87	0.88	0.66	-0.39	1.00	-0.70	-0.55	-0.58	0.30	0.43
qsec	0.74	-0.59	-0.71	-0.66	0.11	-0.70	1.00	0.74	0.59	-0.17	-0.55
vs	0.66	-0.81	-0.71	-0.55	0.13	-0.55	0.74	1.00	0.23	0.36	-0.18
am	0.60	-0.52	-0.49	-0.68	0.79	-0.58	0.59	0.23	1.00	0.43	-0.26
gear	0.48	0.36	0.38	0.43	0.29	0.30	-0.17	0.36	0.43	1.00	-0.12
carb	-0.55	0.55	0.66	0.66	-0.09	0.43	-0.55	-0.18	-0.26	-0.12	1.00

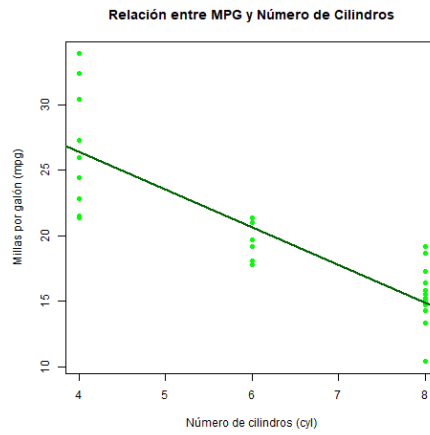
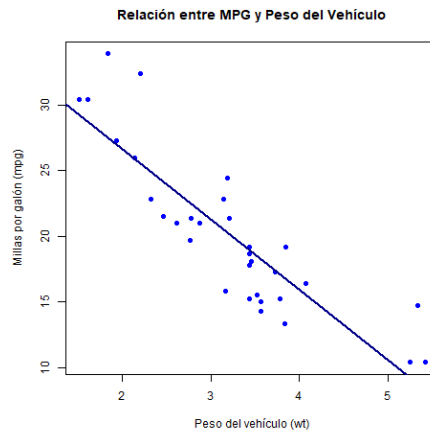
Table 1: Matriz de Correlación de las Variables del Dataset mtcars

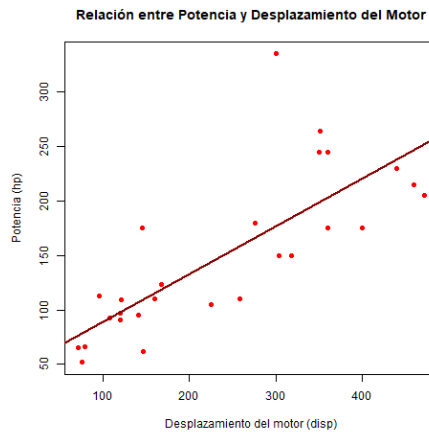
2 Resumen de Hallazgos

- **Correlaciones Fuertes:**

- **mpg y wt:** Correlación negativa fuerte (-0.87), indicando que a mayor peso del automóvil, menor es el rendimiento de combustible.
- **mpg y cyl:** Correlación negativa fuerte (-0.85), sugiriendo que los automóviles con más cilindros tienden a tener un menor rendimiento de combustible.
- **hp y disp:** Correlación positiva fuerte (0.91), indicando que a mayor desplazamiento del motor, mayor es la potencia en caballos de fuerza.

Lo que también es visible en los siguientes gráficos de dispersión que relacionan ambas variables





- **Correlaciones Moderadas:**

- **mpg y drat:** Correlación positiva moderada (0.68), sugiriendo que una mayor relación de transmisión puede estar asociada con un mejor rendimiento de combustible.
- **am y mpg:** Correlación positiva moderada (0.60) entre el tipo de transmisión (manual) y el rendimiento de combustible.

- **Correlaciones Débiles:**

- **gear y mpg:** Correlación moderada (0.48), indicando que el número de marchas tiene un efecto, pero no tan fuerte como otras variables. La correlación moderada sugiere que el número de marchas puede influir en el rendimiento de combustible, pero no es el único factor. Es decir, un automóvil con más marchas podría tener un mejor rendimiento de combustible, pero hay otras variables (como el peso del automóvil, el tipo de motor, etc.) que también juegan un papel importante.
- **drat y carb:** Correlación débil (-0.09), sugiriendo que no hay una relación significativa entre la relación de transmisión y el número de carburadores.

3 Conclusiones

El análisis de la matriz de correlación del conjunto de datos **mtcars** revela varias relaciones significativas entre las variables. Las correlaciones negativas entre el peso y el rendimiento de combustible, así como entre el número de cilindros y el rendimiento de combustible, son particularmente notables. Esto sugiere que los automóviles más pesados y aquellos con más cilindros tienden a ser menos eficientes en términos de consumo de combustible. Por otro lado, las

correlaciones positivas entre el desplazamiento y la potencia sugieren que los motores más grandes tienden a ser más potentes.

Estos hallazgos son útiles para comprender cómo las características de los automóviles se relacionan entre sí y pueden guiar decisiones sobre el diseño y la selección de vehículos en función de su rendimiento y eficiencia.

4 Inserción de Imagen

A continuación se muestra una imagen:

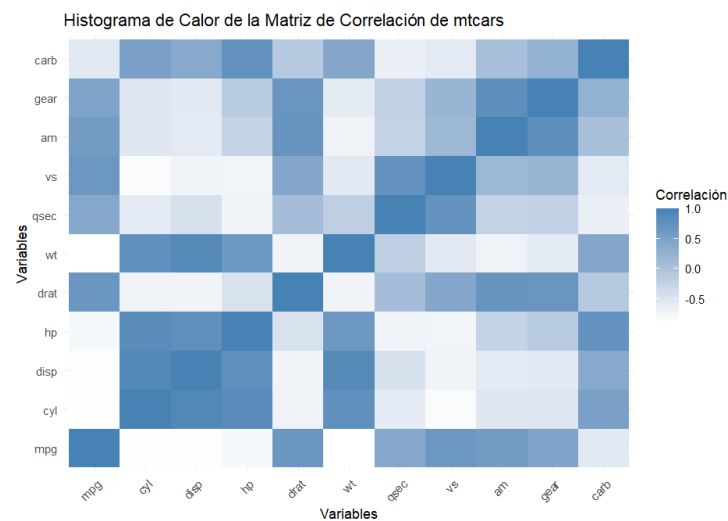


Figure 23: Grafico de Calor , Correlaciones

5 Analisis de la distribucion de Chi-Cuadrado

En el dataset `mtcars`, las variables categóricas que pueden ser analizadas con la prueba Chi-Cuadrado son:

- **vs**: Tipo de motor (0 = V-shaped, 1 = straight).
- **am**: Tipo de transmisión (0 = automática, 1 = manual).
- **gear**: Número de engranajes (3, 4, 5).
- **carb**: Número de carburadores (1, 2, 3, 4, 6, 8).
- **cyl**: Número de cilindros (4, 6, 8).

Variables: `vs` y `am`

- $X^2 = 0.3475355$, $df = 1$, $p\text{-value} = 0.5555115$
- **Conclusión:** No hay evidencia de una relación significativa entre **vs** y **am**

Variables: **vs** y **gear**

- $X^2 = 12.22434$, $df = 2$, $p\text{-value} = 0.002215739$
- **Conclusión:** Hay una relación significativa entre **vs** y **gear**

Variables: **vs** y **carb**

- $X^2 = 15.33968$, $df = 5$, $p\text{-value} = 0.009005379$
- **Conclusión:** Hay una relación significativa entre **vs** y **carb**

Variables: **vs** y **cyl**

- $X^2 = 21.33993$, $df = 2$, $p\text{-value} = 2.323 \times 10^{-5}$
- **Conclusión:** Hay una relación significativa entre **vs** y **cyl**.

Variables: **am** y **gear**

- $X^2 = 20.94467$, $df = 2$, $p\text{-value} = 2.830889 \times 10^{-5}$
- **Conclusión:** Hay una relación significativa entre **am** y **gear**.

Variables: **am** y **carb**

- $X^2 = 6.237131$, $df = 5$, $p\text{-value} = 0.2838241$
- **Conclusión:** No hay evidencia de una relación significativa entre **am** y **carb**

Variables: **am** y **cyl**

- $X^2 = 8.740733$, $df = 2$, $p\text{-value} = 0.01264661$
- **Conclusión:** Hay una relación significativa entre **am** y **cyl**

Variables: **gear** y **carb**

- $X^2 = 16.5181$, $df = 10$, $p\text{-value} = 0.08573092$
- **Conclusión:** No hay evidencia de una relación significativa entre **gear** y **carb**

Variables: **gear** y **cyl**

- $X^2 = 18.03636$, $df = 4$, $p\text{-value} = 0.001214066$
- **Conclusión:** Hay una relación significativa entre **gear** y **cyl**

Variables: `carb` y `cyl`

- $X^2 = 24.38887$, $df = 10$, $p\text{-value} = 0.006632478$
- **Conclusión:** Hay una relación significativa entre `carb` y `cyl`

Este análisis es útil para entender cómo las diferentes características de los automóviles en el dataset `mtcars` están asociadas entre sí. Variables como el tipo de motor (`vs`), el tipo de transmisión (`am`), y el número de cilindros (`cyl`) muestran dependencias importantes que podrían ser exploradas más a fondo en estudios adicionales.

6 Descripción del Gráfico

6.1 Ejes del Gráfico

- **Eje X (Tipo de Transmisión):** Este eje representa el tipo de transmisión de los automóviles, donde:
 - 0 representa transmisión automática.
 - 1 representa transmisión manual.
- **Eje Y (Frecuencia):** Este eje muestra la frecuencia observada de automóviles para cada combinación de tipo de transmisión y tipo de motor.

6.2 Barras

- Cada barra representa la frecuencia de automóviles que tienen una combinación específica de tipo de transmisión y tipo de motor.
- Las barras están agrupadas por el tipo de transmisión (automática y manual) y se diferencian por color según el tipo de motor:
 - 0 representa un motor en línea.
 - 1 representa un motor en V.

6.3 Colores

Los colores de las barras indican el tipo de motor. Esto permite una comparación visual rápida entre las frecuencias de los diferentes tipos de motores para cada tipo de transmisión.

7 Análisis del Gráfico

7.1 Comparación de Frecuencias

Al observar las alturas de las barras, puedes comparar cuántos automóviles tienen transmisión automática frente a manual para cada tipo de motor. Por ejemplo, si la barra para transmisión manual y motor en línea es más alta que la de transmisión automática y motor en línea, esto indica que hay más automóviles con transmisión manual que automática en esa categoría.

7.2 Tendencias Observadas

Puedes identificar tendencias en la distribución de los tipos de transmisión y motor. Por ejemplo:

- Si hay más automóviles con transmisión manual que automática para ambos tipos de motor, esto podría sugerir una preferencia por la transmisión manual en el conjunto de datos.
- Si la frecuencia de motores en línea es mayor en un tipo de transmisión en comparación con el otro, esto puede indicar una relación entre el tipo de transmisión y el tipo de motor.

7.3 Relación entre Variables

El gráfico permite evaluar visualmente si existe una relación entre el tipo de transmisión y el tipo de motor. Si las frecuencias son muy diferentes entre las combinaciones, esto sugiere que podría haber una asociación significativa entre estas variables. Si las frecuencias son similares, podría indicar que el tipo de transmisión no está relacionado con el tipo de motor.

8 Conclusion Distribucion Chi-Cuadrado

- El gráfico de barras es una herramienta poderosa para visualizar la relación entre dos variables categóricas.
- Permite identificar patrones, tendencias y posibles asociaciones entre las variables.
- Junto con los resultados de la prueba chi-cuadrado, este gráfico proporciona una comprensión más profunda de cómo se distribuyen las frecuencias de los automóviles según su tipo de transmisión y motor.

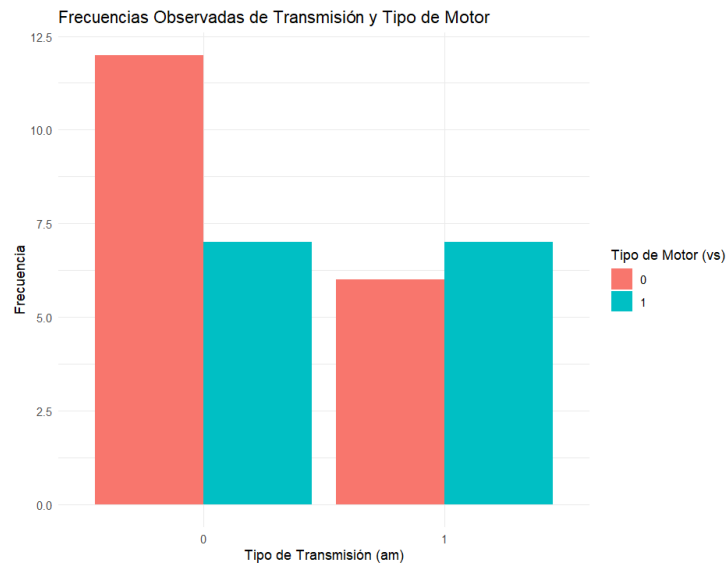


Figure 24: Grafico de Barras, Distribucion Chi-Cuadrado

9 Análisis del Histograma de Distribucion Normalmpg

9.1 Forma de la Distribución

El histograma de `mpg` muestra una distribución que generalmente se asemeja a una campana, lo que sugiere que la mayoría de los automóviles tienen un rendimiento de combustible cercano a la media. La presencia de una forma de campana indica que los datos pueden aproximarse a una distribución normal, aunque es importante verificar esto con pruebas estadísticas adicionales.

9.2 Frecuencia de Valores

Las barras del histograma indican la frecuencia de automóviles en diferentes intervalos de `mpg`. Por ejemplo, si la barra correspondiente al intervalo de 15-20 `mpg` es más alta, esto significa que hay más automóviles en ese rango de eficiencia de combustible. Además, se puede observar cuántos automóviles tienen un `mpg` bajo (por ejemplo, menos de 15) en comparación con aquellos que tienen un `mpg` alto (por ejemplo, más de 30).

9.3 Media y Desviación Estándar

La media de `mpg` puede ser visualmente estimada en el histograma. Si la mayoría de las barras se agrupan alrededor de un valor central, esto indica que la media

se encuentra en ese rango. Por otro lado, la desviación estándar se puede inferir observando la dispersión de las barras: un rango más estrecho de barras indica una menor variabilidad, mientras que un rango más amplio sugiere mayor variabilidad en el rendimiento de combustible. Una desviación estándar de 6.03 para datos que varían entre 10 y 33 indica una alta dispersión relativa. Esto puede ser relevante en contextos donde se espera que los datos sean más homogéneos.

9.4 Identificación de Valores Atípicos

El histograma puede ayudar a identificar valores atípicos (outliers). Si hay barras que se extienden significativamente más allá de la mayoría de las otras, esto podría indicar la presencia de automóviles con un rendimiento de combustible inusualmente bajo o alto. Por ejemplo, si hay un automóvil que tiene un `mpg` muy bajo (por debajo de 10), esto podría ser un valor atípico que merece más investigación.

9.5 Comparación de Rangos de mpg

Al observar el histograma, se puede comparar la frecuencia de automóviles en diferentes rangos de `mpg`. Esto ayuda a identificar qué rangos son más comunes y cuáles son menos frecuentes. Por ejemplo, si notas que hay pocos automóviles con un `mpg` superior a 30, esto sugiere que los automóviles más eficientes en combustible son menos comunes en el conjunto de datos.

9.6 Tendencias Generales

A través del histograma, se pueden identificar tendencias generales en el rendimiento de combustible de los automóviles en el conjunto de datos. Esto puede ser útil para entender el mercado automotriz o para realizar análisis de eficiencia energética. Si la mayoría de los automóviles tienen un `mpg` en el rango de 20-25, esto podría indicar que los vehículos en este conjunto de datos son relativamente eficientes en comparación con los estándares de consumo de combustible.

9.7 Conclusión de la Distribución Normal

El análisis del histograma de `mpg` proporciona información valiosa sobre la distribución del rendimiento de combustible en el conjunto de datos `mtcars`. Permite visualizar la forma de la distribución, identificar frecuencias, evaluar la media y la variabilidad, detectar valores atípicos y comprender las tendencias generales en el rendimiento de los automóviles. Este análisis es fundamental para realizar inferencias estadísticas y tomar decisiones informadas sobre el rendimiento de combustible de los vehículos.

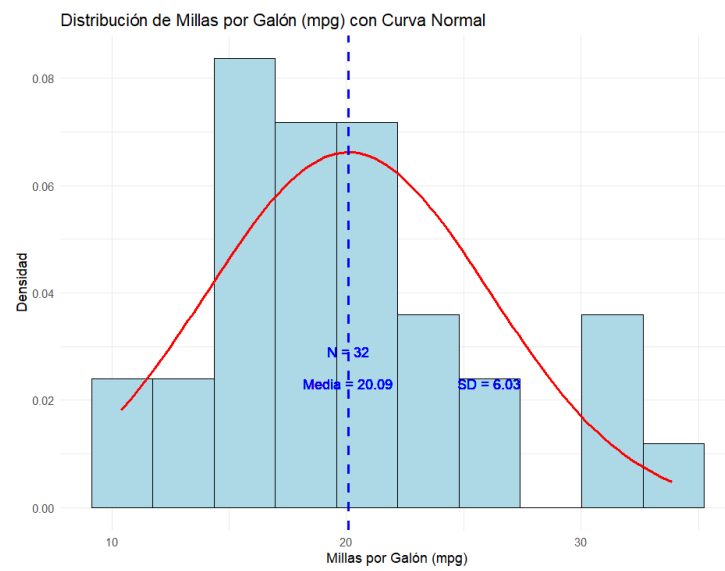


Figure 25: Distribucion Normal