Universidad Nacional de Colombia
Departamento de Estadística
Análisis de regresión (II-2022)
Parcial II

Andrés Mauricio Rico Parada - aricop@unal.edu.co
Juan David Carrrascal Ibañez - jdcarrascali@unal.edu.co

1. a)

$$\begin{cases} Y_k = \mu_k + e_k \\ \mu_k = \beta \\ e_k \sim N(0, \sigma^2) \\ e_1, e_2, ..., e_k \quad \text{independientes} \end{cases}$$

Dado que los errores son independientes y todos se distribuyen con una distribución una normal estándar de la forma $N(0, \sigma^2)$, tenemos que el modelo se puede escribir de forma matricial como:

$$\begin{cases} \mathbb{Y} = \mu + \mathbb{E} \\ \mu = \mathbf{1}\beta \\ \mathbb{E} \sim MVN(0, \sigma^2 \mathbb{I}) \end{cases}$$

b)

$$\begin{aligned} \hat{\beta} &= argmin(Q(\beta)) \\ &= argmin(\sum_{i=1}^{n} (y_i - \beta)^2) \\ &= argmin((\mathbb{Y} - \mathbf{1}\beta)^T (\mathbb{Y} - \mathbf{1}\beta)) \\ &= argmin(\mathbb{Y}^T \mathbb{Y} - \mathbb{Y}\mathbf{1}\beta - \beta \mathbf{1}^T \mathbb{Y} + \beta^2 \mathbf{1}^T \mathbf{1}) \\ &= argmin(\mathbb{Y}^T \mathbb{Y} - 2\mathbb{Y}^T \mathbf{1}\beta + n\beta^2) \end{aligned}$$

Derivando e igualando a cero:

$$\frac{dQ}{d\beta} = -2 \cdot \mathbf{1}^T \mathbb{Y} + 2n\beta = 0$$
$$\therefore \hat{\beta} = \frac{1}{n} \mathbf{1}^T \mathbb{Y} = \bar{Y}$$

Donde $\bar{Y}$ representa la media muestral de $\{Y_k\}$. Ahora bien, como

$$\frac{d^2Q}{d\beta} = 2n > 0$$

Entonces $\hat{\beta}$ es un mínimo local y dado que $2n$ no depende de β entonces $Q(\beta)$ es una función convexa y eso hace que la solución sea un mínimo global. Además completamos la prueba garantizando que:

$$\lim_{||\beta|| \to \infty} Q(\beta) = \infty$$

c) Varianza de $\hat{\beta}$

$$Var(\hat{\beta}) = Var(\frac{1}{n}\mathbf{1}^T\mathbb{Y}) = \frac{1}{n^2}Var(\mathbf{1}^T\mathbb{Y}) = \frac{1}{n^2}\mathbf{1}^T Var(Y)\mathbf{1} = \sigma^2\frac{1}{n^2}\mathbf{1}^T\mathbf{1} = \frac{\sigma^2}{n}$$

Este resultado corresponde a la varianza de la media muestral $\bar{Y}_n$.

d) Se tiene que $E(Y_k) = \beta$ y $Var(Y_k) = \sigma^2$, para $k = 1, ..., n$. Además, debido a los errores, que tienen una distribución normal y son independientes, se tiene que $\{Y_k\}$ es una muestra aleatoria y todos las propiedades de la media muestral aplican para $\hat{\beta}$. Por lo tanto, $\hat{\beta}$ además de ser insesgado, es consistente y es un $UMVUE$.

2.

a. Recordemos que $\bar{\mathbb{Y}}_n = \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{Y}$ y $\widehat{\bar{\mathbb{Y}}_n} = \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{H}\mathbb{Y}$. De manera tal que:

$$\begin{aligned}
\sum_{i=1}^{n}(Y_i - \bar{Y}_n)(\widehat{Y}_i - \bar{\hat{Y}}_n) &= (\mathbb{Y} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{Y})^T(\mathbb{H}\mathbb{Y} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{H}\mathbb{Y}) \\
&= (\mathbb{Y}^T - \frac{1}{n}\mathbb{Y}^T\mathbf{1}\mathbf{1}^T)(\mathbb{H} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{H})\mathbb{Y} \\
&= \mathbb{Y}^T(\mathbb{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)(\mathbb{H} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\mathbb{Y} \\
&= \mathbb{Y}^T(\mathbb{H} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{H} - \frac{1}{n}\mathbf{1}\mathbf{1}^T + \frac{1}{n^2}\mathbf{1}\mathbf{1}^T\mathbf{1}\mathbf{1}^T)\mathbb{Y} \\
&= \mathbb{Y}^T(\mathbb{H} - \frac{2}{n}\mathbf{1}\mathbf{1}^T\mathbb{H} + \frac{1}{n^2}\mathbf{1}n\mathbf{1}^T)\mathbb{Y} \\
&= \mathbb{Y}^T(\mathbb{H} - \frac{2}{n}\mathbf{1}\mathbf{1}^T + \frac{1}{n}\mathbf{1}\mathbf{1}^T)\mathbb{Y} \\
&= \mathbb{Y}^T(\mathbb{H} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\mathbb{Y} = SC_{mod}
\end{aligned}$$

Teniendo en cuenta que:

$$\widehat{\bar{\mathbb{Y}}}_n = \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{H}\mathbb{Y} = \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbb{Y} = \bar{\mathbb{Y}}_n$$

Concluimos que:

$$\begin{aligned} r &= \frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)(\widehat{Y}_i - \bar{\widehat{Y}}_n)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \sum_{i=1}^n (\widehat{Y}_i - \bar{\widehat{Y}}_n)^2}} \\ r &= \frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)(\widehat{Y}_i - \bar{\widehat{Y}}_n)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \sum_{i=1}^n (\widehat{Y}_i - \bar{Y}_n)^2}} \\ &= \frac{SC_{mod}}{\sqrt{SC_{Tot} SC_{mod}}} = \sqrt{\frac{SC_{mod}}{SC_{Tot}}} \\ &\therefore r^2 = R^2 \end{aligned}$$

b. Sea $k \in \{1, 2, ..., p\}$. Luego,

- $$\begin{aligned} \sum_{i=1}^n r_i \hat{Y}_i &= \sum_{i=1}^n (Y_i - \hat{Y}_i)\hat{Y}_i \\ &= (\mathbb{Y} - \mathbb{H}\mathbb{Y})^T \mathbb{H}\mathbb{Y} \\ &= (\mathbb{Y}^T - \mathbb{Y}^T\mathbb{H})\mathbb{H}\mathbb{Y} \\ &= \mathbb{Y}^T(\mathbb{I} - \mathbb{H})\mathbb{H}\mathbb{Y} \\ &= 0\mathbb{Y}^T\mathbb{Y} = 0 \end{aligned}$$

- $$\begin{aligned} \sum_{i=1}^n r_i x_{ki} &= \sum_{i=1}^n (Y_i - \hat{Y}_i) x_{ki} \\ &= (\mathbb{Y} - \mathbb{H}\mathbb{Y})^T \mathbb{X} \\ &= (\mathbb{Y}^T - \mathbb{Y}^T\mathbb{H})\mathbb{H}\mathbb{X} \\ &= \mathbb{Y}^T(\mathbb{I} - \mathbb{H})\mathbb{H}\mathbb{X} \\ &= 0\mathbb{Y}^T\mathbb{H}\mathbb{X} = 0 \end{aligned}$$

c. La matriz de proyección para el modelo 2 será $\mathbb{H}_2 = \widetilde{\mathbb{X}}(\widetilde{\mathbb{X}}^T\widetilde{\mathbb{X}})^{-1}\widetilde{\mathbb{X}}^T$. De modo que

$$\mathbb{H}_1\mathbb{H}_2 = [\mathbb{H}_1\widetilde{\mathbb{X}}](\widetilde{\mathbb{X}}^T\widetilde{\mathbb{X}})^{-1}\widetilde{\mathbb{X}}^T = \widetilde{\mathbb{X}}(\widetilde{\mathbb{X}}^T\widetilde{\mathbb{X}})^{-1}\widetilde{\mathbb{X}}^T = \mathbb{H}_2$$

Como $\mathbb{H}_1$ y $\mathbb{H}_2$ son simétricas entonces

$$\mathbb{H}_2\mathbb{H}_1 = \mathbb{H}_2^T\mathbb{H}_1^T = (\mathbb{H}_1\mathbb{H}_2)^T = (\mathbb{H}_2)^T = \mathbb{H}_2$$

De esta forma concluimos que $\mathbb{H}_2\mathbb{H}_1 = \mathbb{H}_1\mathbb{H}_2 = \mathbb{H}_2$.

d.Haremos uso del siguiente teorema:

Teorema 18 *Si* $\mathbf{Y} \sim MVN(\mu_{\mathbf{Y}}, \mathbf{\Sigma_Y})$ *con* $\mathbf{\Sigma_Y}$ *no singular,* $\mathbf{Y}^T\mathbf{AY} \sim \chi^2(r(\mathbf{A}), \mu_{\mathbf{Y}}{}^T\mathbf{A}\mu_{\mathbf{Y}})$ *si y solo si* $\mathbf{A\Sigma_Y}$ *es idempotente, con* $\mathbf{A}$ *una matriz de dimensiones* $n \times n$.

Notemos que en el modelo 1, $\mathbb{Y} \sim MVN(\mathbb{X}\beta, \sigma^2\mathbb{I})$ y $\sigma^2\mathbb{I}$ es no singular pues es definida positiva. Como $\frac{(\mathbb{H}_2-\mathbb{H}_1)}{\sigma^2}\sigma^2\mathbb{I} = (\mathbb{H}_2 - \mathbb{H}_1) = (\mathbb{H}_2 - \mathbb{H}_1)^2$ entonces $\frac{(\mathbb{H}_2-\mathbb{H}_1)}{\sigma^2}\sigma^2\mathbb{I}$ es idempotente.

Ahora bien, sea $k := p - p_2$ y notemos que

$$r(\frac{\mathbb{H}_2 - \mathbb{H}_1}{\sigma^2}) = r(\mathbb{H}_2 - \mathbb{H}_1) = tr(\mathbb{H}_2 - \mathbb{H}_1) = tr(\mathbb{H}_2) - tr(\mathbb{H}_1) = p - p_2 = k$$

Por otra parte,

$$\begin{aligned}
\mathbb{Y}^T(\mathbb{H}_1 - \mathbb{H}_2)\mathbb{Y} &= \mathbb{Y}^T(\mathbb{H}_2 - \mathbb{H}_1)(\mathbb{H}_2 - \mathbb{H}_1)\mathbb{Y} \\
&= (\mathbb{Y}^T\mathbb{H}_2 - \mathbb{Y}^T\mathbb{H}_1)(\mathbb{H}_2\mathbb{Y} - \mathbb{H}_1\mathbb{Y}) \\
&= (\widehat{\mathbb{Y}}_2^T - \widehat{\mathbb{Y}}_1^T)(\widehat{\mathbb{Y}}_2 - \widehat{\mathbb{Y}}_1) \\
&= (\widehat{\mathbb{Y}}_2 - \widehat{\mathbb{Y}}_1)^T(\widehat{\mathbb{Y}}_2 - \widehat{\mathbb{Y}}_1) \\
&= ||\widehat{\mathbb{Y}}_2 - \widehat{\mathbb{Y}}_1||^2
\end{aligned}$$

Con estas tres condiciones garantizadas podemos hacer uso del teorema 18 para construir una función con distribución Chi-cuadrado con k grados de libertad como se indica a continuación:

$$\frac{||\widehat{\mathbb{Y}}_2 - \widehat{\mathbb{Y}}_1||^2}{\sigma^2} = \frac{\mathbb{Y}^T(\mathbb{H}_1 - \mathbb{H}_2)\mathbb{Y}}{\sigma^2} \sim \chi^2(k)$$

Para construir la distribución F requerida, necesitamos hacer uso del siguiente teorema:

Teorema 20 *Sea* $\mathbf{Y} \sim MVN(\mu_{\mathbf{Y}}, \mathbf{\Sigma_Y})$ *con* $\mathbf{\Sigma_Y}$ *no singular.* $\mathbf{Y}^T\mathbf{AY}$ *y* $\mathbf{Y}^T\mathbf{BY}$ *son estadísticamente independientes si* $\mathbf{B\Sigma_Y A} = \mathbf{O}$*, con* $\mathbf{A}$ *y* $\mathbf{B}$ *matrices de dimensiones* $n \times n$.

De nuevo, usamos que $\mathbb{Y} \sim MVN(\mathbb{X}\beta, \sigma^2\mathbb{I})$. Para usar la notación del teorema consideremos:

$$\begin{aligned}
\frac{\mathbb{Y}^T(\mathbb{H}_1 - \mathbb{H}_2)\mathbb{Y}}{\sigma^2} &= \mathbb{Y}^T A\mathbb{Y} \\
\frac{\mathbb{Y}^T(\mathbb{I} - \mathbb{H}_1)\mathbb{Y}}{\sigma^2} &= \mathbb{Y}^T B\mathbb{Y}
\end{aligned}$$

Así, como

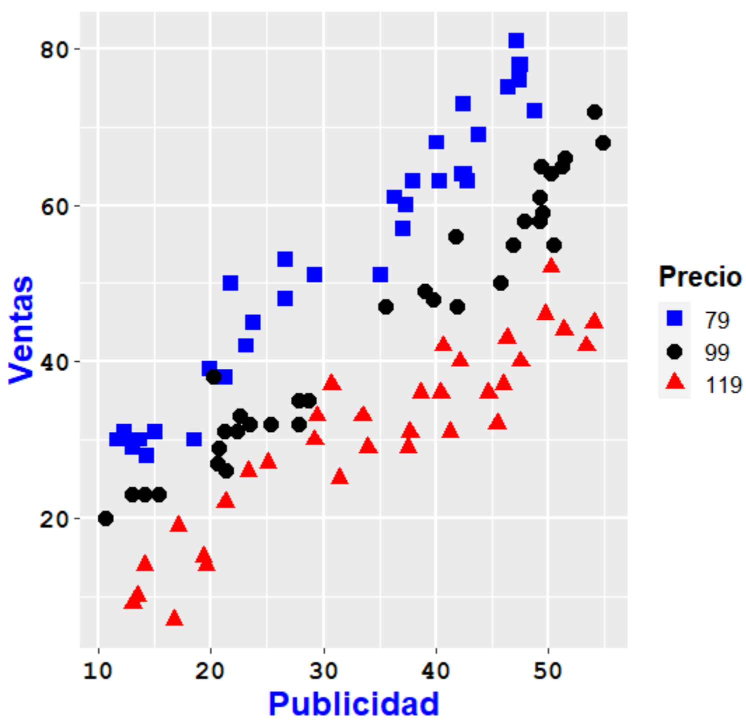
$$(\mathbb{H}_1 - \mathbb{H}_2)(\mathbb{I} - \mathbb{H}_1) = \mathbb{H}_1 - \mathbb{H}_2 - \mathbb{H}_1^2 + \mathbb{H}_2\mathbb{H}_1 = \mathbb{H}_1 - \mathbb{H}_2 - \mathbb{H}_1 + \mathbb{H}_2 = 0$$

Concluimos que $\frac{\mathbb{Y}^T(\mathbb{H}_1-\mathbb{H}_2)\mathbb{Y}}{\sigma^2}$ y $\frac{\mathbb{Y}^T\mathbb{I}-\mathbb{H}_1\mathbb{Y}}{\sigma^2}$ son estadísticamente independientes. Además, recor-

demos que $\hat{\sigma}^2 = \frac{\mathbb{Y}^T(\mathbb{I}-\mathbb{H}_1)\mathbb{Y}}{n-p}$ que cumple $\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p)$.
Finalmente estamos en condiciones de grantizar que:

$$\frac{\frac{||\widehat{\mathbb{Y}}_2-\widehat{\mathbb{Y}}_1||^2}{\sigma^2} \Big/ k}{\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \Big/ n-p} = \frac{||\widehat{\mathbb{Y}}_2-\widehat{\mathbb{Y}}_1||^2}{k\hat{\sigma}^2} \sim F(k, n-p)$$

1) a)



Se puede observar en el gráfico que hay una relación lineal entre las ventas y la publicidad y también se ve una clara diferencia en las ventas de acuerdo al precio. Por lo tanto, estas dos variables parecen ser relevantes para explicar las ventas, pero además, se agregará la interacción de estas dos variables al modelo ya que el efecto del aumento de la publicidad puede que dependa de los diferentes precios (como se puede ver en el gráfico del punto b).

b)

$$\begin{cases} Y_k = \mu_k + e_k \\ \mu_k = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 \\ e_k \sim N(0, \sigma^2) \\ e_1, e_2, ..., e_k \quad \text{independientes} \end{cases}$$

Dónde x_1 es la inversión en publicidad (dólares), las variables x_2 y x_3 corresponden a

$$x_2 = \begin{cases} 1, \text{Precio 99 centavos} \\ 0, \quad \text{e.o.c.} \end{cases} \qquad x_3 = \begin{cases} 1, \text{Precio 119 centavos} \\ 0, \quad \text{e.o.c.} \end{cases}$$

c)

```
minimercados2 <- dummy_cols(minimercados, c("precio"),remove_first_dummy =
    TRUE,remove_selected_columns = TRUE)
```

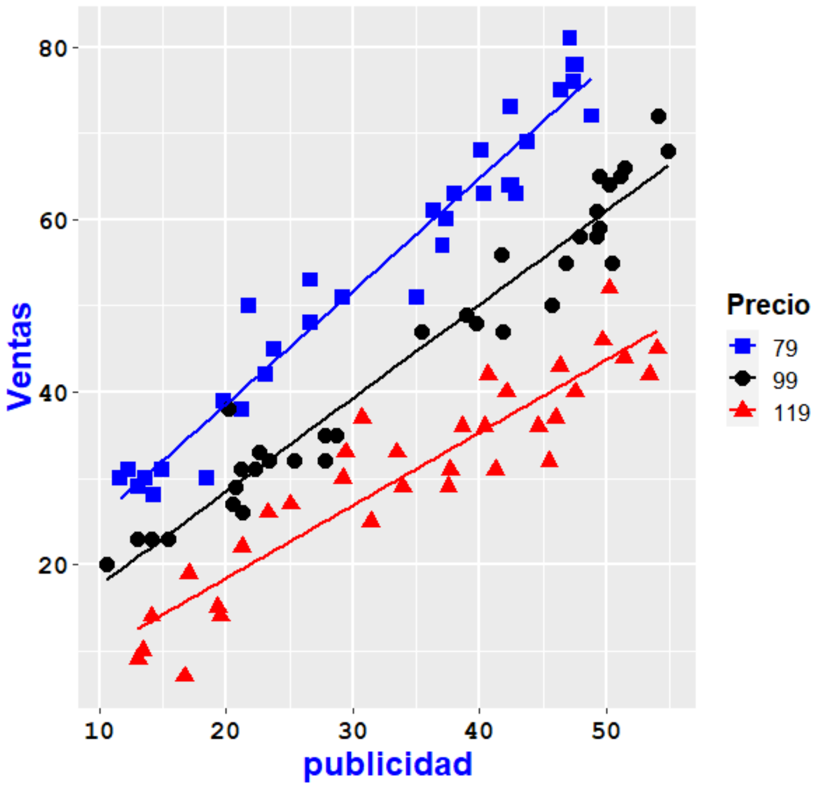
```
Call:
lm(formula = ventas ~ 1 + publicidad + precio_99 + precio_119 +
    publicidad * precio_99 + publicidad * precio_119, data = minimercados2
   )


Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            12.39153    1.90922   6.490 3.99e-09 ***
publicidad              1.31126    0.05567  23.554  < 2e-16 ***
precio_99              -5.64112    2.64231  -2.135 0.035369 *
precio_119            -10.83863    2.78344  -3.894 0.000184 ***
publicidad:precio_99   -0.22636    0.07394  -3.061 0.002874 **
publicidad:precio_119  -0.46728    0.07836  -5.963 4.31e-08 ***
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1
           1

Residual standard error: 3.994 on 94 degrees of freedom
Multiple R-squared:  0.9521,  Adjusted R-squared:  0.9495
F-statistic: 373.6 on 5 and 94 DF,  p-value: < 2.2e-16
```

Interpretación:

- $\hat{\beta}_0$: Promedio de ventas de máquinas de afeitar de 79 centavos cuando no hay inversión en publicidad.
- $\hat{\beta}_1$: Aumento promedio de ventas de máquinas de afeitar de 79 centavos por cada dólar de inversión en publicidad.
- $\hat{\beta}_2$: Decremento promedio de ventas de máquinas de afeitar de 99 centavos con respecto a las de 79 centavos cuando no hay inversión en publicidad.
- $\hat{\beta}_3$: Decremento promedio de ventas de máquinas de afeitar de 99 centavos con respecto a las de 79 centavos cuando no hay inversión en publicidad.
- $\hat{\beta}_4$ Decremento promedio de ventas de máquinas de afeitar de 99 centavos con respecto al aumento promedio de las de 79 centavos por cada dólar de inversión en publicidad.

- $\hat{\beta}_5$: Decremento promedio de ventas de máquinas de afeitar de 119 centavos con respecto al aumento promedio de las de 79 centavos por cada dólar de inversión en publicidad.

- $\hat{\sigma}$: Distancia promedio del valor predicho por el modelo al valor observado.



d) Se usará el comando ANOVA para ver la significancia de la variable precio en el modelo. Para usarlo hay que validar primero la homoscedasticidad.

```
studentized Breusch-Pagan test
data:  fit
BP = 3.3094, df = 5, p-value = 0.6524
```

Ahora se hará el modelo fit2 sin esta variable.

```
fit2 <- lm(ventas~1+publicidad,minimercados2)
anova(fit2,fit,test="F")

Analysis of Variance Table

Model 1: ventas ~ 1 + publicidad
Model 2: ventas ~ 1 + publicidad + precio_99 + precio_119 + publicidad *
    precio_99 + publicidad * precio_119
  Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1     98 13468.9
2     94  1499.8  4     11969 187.54 < 2.2e-16 ***
```

Con este p-valor se rechaza la hipótesis nula, por lo tanto, hay evidencia suficiente para decir que el precio es significativo sobre la cantidad promedio de paquetes vendidos de máquinas de afeitar.

e) Para este pregunta se juzgará el sistema de hipótesis

$$\begin{cases} H_0 : \beta_4 = 0, \beta_5 = 0 \\ H_1 : \beta_4 \neq 0 \text{ o } \beta_5 \neq 0 \end{cases}$$

Se puede realizar con el comando linearHypothesis, juzgando una prueba de la forma

$$\begin{cases} H_0 : \mathbf{C}\beta = 0_{6x1} \\ H_1 : \mathbf{C}\beta \neq 0_{6x1} \end{cases}$$

con $C = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

```
linearHypothesis(fit,c("publicidad:precio_119=0","publicidad:precio_99=0"))

Hypothesis:
publicidad:precio_119 = 0
publicidad:precio_99 = 0

Model 1: restricted model
Model 2: ventas ~ 1 + publicidad + precio_99 + precio_119 + publicidad *
    precio_99 + publicidad * precio_119

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     96 2067.6
2     94 1499.8  2    567.74 17.791 2.801e-07 ***
```

Con este p-valor se rechaza la hipótesis nula, por lo tanto, hay evidencia suficiente para decir que el efecto de la inversión en publicidad sobre el número promedio de paquetes vendidos de máquinas de afeitar depende del precio.

f) Como el efecto de la inversión en publicidad para el precio de 79 centavos esta dado por β_1, si el efecto es igual esto es equivalente a que $\beta_5 = 0$, así que para este punto hay que juzgar el sistema de hipótesis

$$\begin{cases} H_0 : \beta_5 = 0 \\ H_1 : \beta_5 \neq 0 \end{cases}$$

La salida de summary(fit) (punto c), ya juzga este sistema de hipótesis, dando un p-valor de $4{,}31e-08$, por lo tanto, hay evidencia suficiente para decir que el efecto de la inversión en publicidad sobre el número promedio de paquetes vendidos es diferente para los minimercados que venden las máquinas de afeitar a 119 centavos que para aquellos que venden 79 centavos.

g) Como el número promedio de ventas cuando no se invierte en publicidad en los minimercados que venden a 79 centavos es β_0 y el número promedio de ventas cuando no se invierte en publicidad en los minimercados que venden a 99 centavos es $\beta_0+\beta_2$, hay que juzgar el sistema de hipótesis

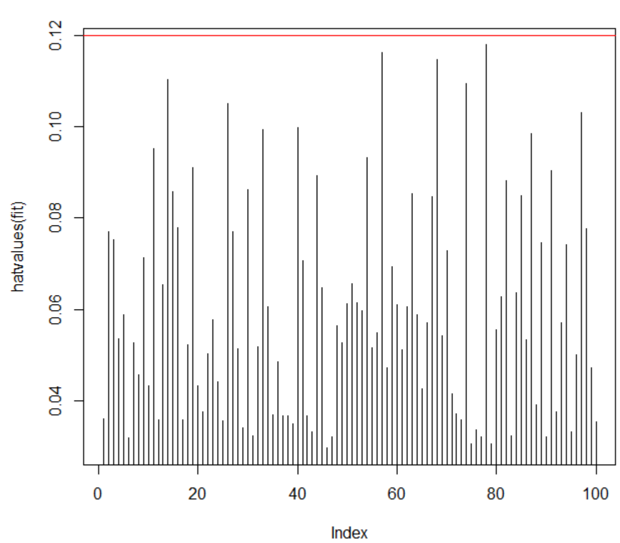
$$\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$$

La salida de summary(fit) (punto c), ya juzga este sistema de hipótesis, dando un p-valor de 0,035369, por lo tanto, con una confianza de 95 % hay evidencia suficiente para decir que cuando no se invierte en publicidad, el número promedio de paquetes vendidos es diferente en los minimercados que venden las máquinas de afeitar a 99 centavos que en aquellos que venden a 79 centavos.

h) Observaciones de Alto leverage

Como se puede observar en el gráfico, en este caso, parece no haber observaciones de alto leverage.

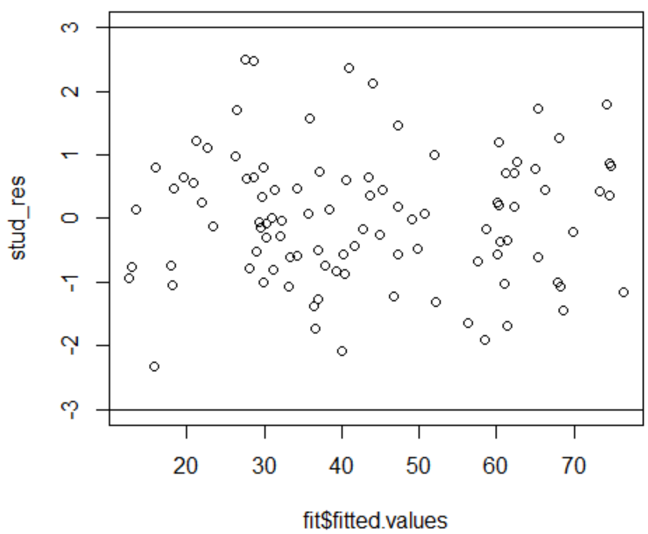
```
n<-dim(minimercados2[,2])[1]
plot(hatvalues(fit),type="h")
abline(h=2*6/n,col="red")
```



Observaciones atípicas

Como se puede observar en el gráfico, parece no haber observaciones atípicas, ya que ninguno de los residuales studentizados es mayor a 3.

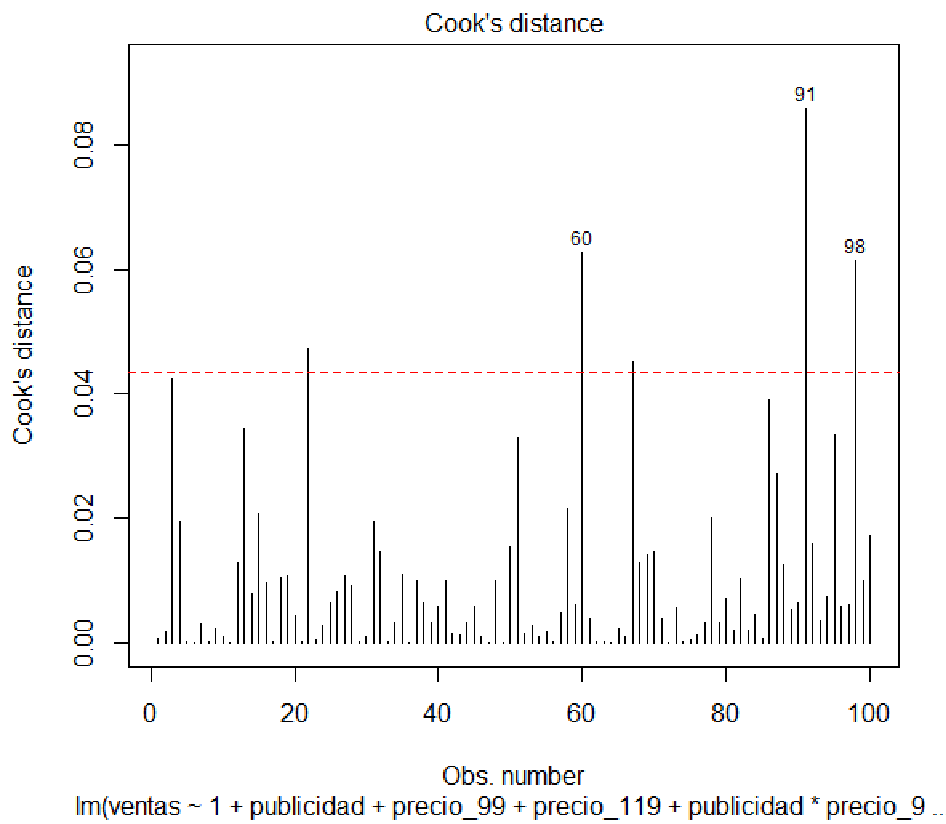
```
stud_res<-studres(fit)
plot(fit$fitted.values,stud_res,ylim=c(-3,3))
abline(h=3)
abline(h=-3)
```



Observaciones Influyentes

De acuerdo a la distancia de Cook, las observaciones 22,60,67,91 y 98 pueden ser influyentes.

```
corte <- 4/(n-length(fit$coefficients)-2)
plot(fit, which=4, cook.levels=corte)
abline(h=corte, lty=2, col="red")
cooksd<-cooks.distance(fit)
cooksd[which(cooksd>corte)]

        22         60         67         91         98
0.04727864 0.06271481 0.04532197 0.08592108 0.06150164
```



Estimación del modelo sin observaciones influyentes

```
fit3<-update(fit,subset={setdiff(row(minimercados2)[,2], c(22,60,67,91,98))})
summary(fit3)

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            11.50729    1.72924   6.655 2.25e-09 ***
publicidad              1.32994    0.04996  26.620  < 2e-16 ***
precio_99              -5.29253    2.41659  -2.190 0.031133 *
precio_119             -7.86178    2.57036  -3.059 0.002937 **
publicidad:precio_99   -0.24389    0.06740  -3.619 0.000491 ***
publicidad:precio_119  -0.54591    0.07200  -7.582 3.10e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.548 on 89 degrees of freedom
Multiple R-squared:  0.9614,	Adjusted R-squared:  0.9592
F-statistic: 443.3 on 5 and 89 DF,  p-value: < 2.2e-16
```

Se puede observar que removiendo las observaciones influyentes cambia la estimación de los parámetros de localización pero no por mucho, son similares, entonces parece que estas observaciones no están afectanto tanto el modelo. La estimación de σ pasa de 3.994 a 3.548, por lo tanto, en este caso el modelo está explicando un poco mejor el comportamiento de los datos.