```
> summary(modelo) #Obteniendo tabla de parámetros ajustados
Call:
lm(formula = Porc.Pureza02 ~ Porc.Hidrocarb)

Residuals:
     Min       1Q   Median       3Q      Max
-1.83029 -0.73334  0.04497  0.69969  1.96809

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)      74.283      1.593   46.62  < 2e-16 ***
Porc.Hidrocarb   14.947      1.317   11.35 1.23e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.087 on 18 degrees of freedom
Multiple R-squared:  0.8774,    Adjusted R-squared:  0.8706
F-statistic: 128.9 on 1 and 18 DF,  p-value: 1.227e-09
```

Y : % de oxígeno producido ✓

X : % de hidrocarburos ✓

Modelo Estadístico

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

Línea de regresión ajustada

$$\hat{y}_i = 74.283 + 14.947 X_i$$

Por cada unidad que cambie el % de hidrocarburos, el % de oxígeno producido en promedio aumenta en 14.947

Vamos a calcular $E(\hat{\beta}_1)$ y $V(\hat{\beta}_1)$

Tenemos:

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$$

$$c_i = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2} \qquad \therefore \boxed{\hat{\beta}_1 = \sum_{i=1}^{n} c_i y_i}$$

$$\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} \left(\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \left[\sum_{i=1}^{n} (x_i - \bar{x}) \right]$$

$$= \frac{1}{\sum (x_i - \bar{x})^2} \left[\sum_{i=1}^{n} x_i - \underbrace{n\bar{x}}_{\cancel{n} \frac{\sum x_i}{\cancel{n}}} \right] = 0$$

$$\therefore \sum_{i=1}^{n} c_i = 0$$

$$\sum_{i=1}^{n} c_i x_i = \sum_{i=1}^{n} \left[\frac{(x_i - \bar{x})\, x_i}{\sum (x_i - \bar{x})^2} \right]$$

$$= \frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \left[\sum_{i=1}^{n} (x_i - \bar{x}) x_i \right]$$

$$\parallel$$

$$\sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x}) = \sum_{i=1}^{n} (x_i - \bar{x}) x_i - \bar{x} \underbrace{\sum_{i=1}^{n} (x_i - \bar{x})}_{\to 0}$$

$$= \sum_{i=1}^{n} (x_i - \bar{x}) x_i$$

$$\therefore \sum_{i=1}^{n} c_i x_i = \frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sum (x_i - \bar{x})(x_i - \bar{x})$$

$$= 1$$

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} c_i y_i\right)$$

$$= \sum_{i=1}^{n} c_i E(y_i)$$

$$= \sum_{i=1}^{n} c_i (\beta_0 + \beta_1 X_i)$$

$$= \beta_0 \underbrace{\cancel{\sum c_i}}_{0} + \beta_1 \underbrace{\cancel{\sum c_i X_i}}_{1} = \beta_1$$

Por lo tanto $\hat{\beta}_1$ es un estimador insesgado de β_1.

Para calcular la varianza necesitamos el siguiente resultado

$$\sum_{i=1}^{n} c_i^2 = \sum_{i=1}^{n} \left[\frac{x_i - \bar{X}}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \right]^2$$

$$= \frac{\cancel{\sum_{i=1}^{n} (X_i - \bar{X})^2}}{\left[\sum_{i=1}^{n} (x_i - \bar{X})^2\right]^2} = \frac{1}{\sum_{i=1}^{n} (x_i - \bar{X})^2}$$

$$V(\hat{\beta}_1) = V\left(\sum_{i=1}^{n} c_i y_i\right)$$

$$= \underbrace{\sum_{i=1}^{n} c_i^2}\, V(y_i)$$

$$= \frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\, \sigma^2 \rightarrow \text{No es conocido}$$

$$V(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = MSE$$

$$\therefore \quad \hat{\beta}_1 \sim \text{Normal}\left(\beta_1, \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

$$\frac{\hat{\beta}_1 - \beta_1}{\dfrac{\sigma}{\sqrt{\sum(x_i - \bar{x})^2}}} \sim N(0, 1)$$

Como no conocemos σ, entonces tenemos

$$\frac{\hat{\beta}_1 - \beta_1}{\frac{\hat{\sigma}}{\sqrt{\sum (x_i - \bar{x})^2}}} \sim t_{n-2}$$

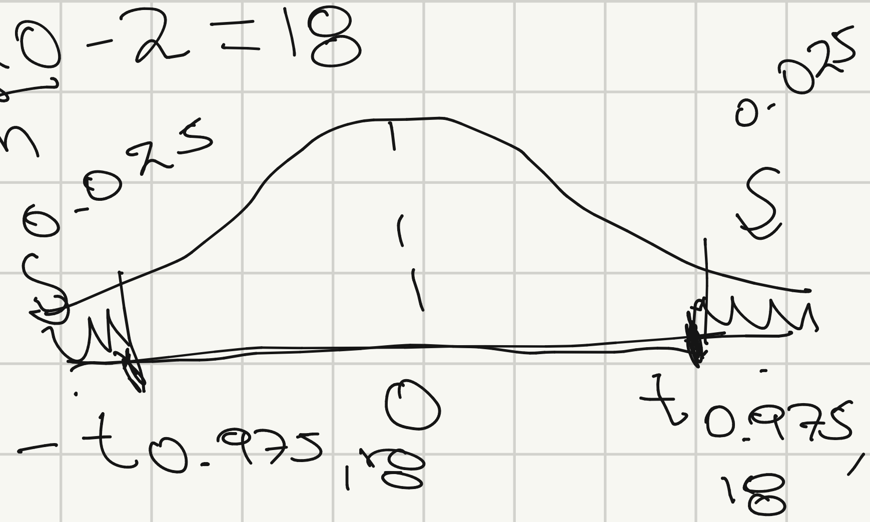
```
> summary(modelo) #Obteniendo tabla de parámetros ajustados
Call:
lm(formula = Porc.PurezaO2 ~ Porc.Hidrocarb)

Residuals:
     Min       1Q   Median       3Q      Max
-1.83029 -0.73334  0.04497  0.69969  1.96809

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)      74.283      1.593   46.62  < 2e-16 ***
Porc.Hidrocarb   14.947      1.317   11.35 1.23e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.087 on 18 degrees of freedom
Multiple R-squared:  0.8774,    Adjusted R-squared:  0.8706
F-statistic: 128.9 on 1 and 18 DF,  p-value: 1.227e-09
```

$d.e.(\hat{\beta}_0) = 1.593$

$d.e.(\hat{\beta}_1) = 1.317$

$\alpha = 5\%$

$\hat{\sigma} = 1.087$ $\quad \underbrace{20}_{n} - 2 = 18$



$H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$

$$t_0 = \frac{\hat{\beta}_1 - \beta_1}{d.e.(\hat{\beta}_1)} = \frac{14.947 - 0}{1.317} = 11.35$$

p-value $\approx 0 < \alpha$ $\therefore$ hay evidencia para rechazar la hipótesis nula.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\beta_1 \neq 0$$

Por lo tanto hay evidencia para decir que hay una relación estadística entre el % de oxígeno y el % de hidrocarburos. Entonces por cada que incremente el % de hidrocarburos, en promedio el % de oxígeno va a aumentar en 14.947.

$$Y_i \mid X_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$$

$$P(Y_i \mid X_i > a)$$

$$P\left(Z_i > \frac{a - (\beta_0 + \beta_1 X_i)}{\sigma}\right)$$