Tarea 2

JESUS ALEXIS SANCHEZ MORENO

MATRICULA: 224470329

Índice de ejercicios

EJERCICIO-37

3.7 We have 2008 data on *INCOME* = income per capita (in thousands of dollars) and *BACHELOR* = percentage of the population with a bachelor's degree or more for the 50 U.S. States plus the District of Columbia, a total of $N = 51$ observations. The results from a simple linear regression of *INCOME* on *BACHELOR* are

$$\begin{array}{rccc} \widehat{INCOME} = & (a) & + & 1.029BACHELOR \\ \text{se} & (2.672) & & (c) \\ t & (4.31) & & (10.75) \end{array}$$

a. Using the information provided calculate the estimated intercept. Show your work.
b. Sketch the estimated relationship. Is it increasing or decreasing? Is it a positive or inverse relationship? Is it increasing or decreasing at a constant rate or is it increasing or decreasing at an increasing rate?
c. Using the information provided calculate the standard error of the slope coefficient. Show your work.
d. What is the value of the t-statistic for the null hypothesis that the intercept parameter equals 10?
e. The p-value for a two-tail test that the intercept parameter equals 10, from part (d), is 0.572. Show the p-value in a sketch. On the sketch, show the rejection region if $\alpha = 0.05$.
f. Construct a 99% interval estimate of the slope. Interpret the interval estimate.
g. Test the null hypothesis that the slope coefficient is one against the alternative that it is not one at the 5% level of significance. State the economic result of the test, in the context of this problem.

a. Cálculo del Intercepto (a)

La relación entre el estadístico t, el coeficiente y su error estándar es:

$$t = \frac{\text{coeficiente}}{\text{se}}$$

Para encontrar el intercepto (a), despejamos el coeficiente:

$$a = t_{\text{intercepto}} \times se_{\text{intercepto}}$$

Sustituyendo los valores dados:

$$a = 4.31 \times 2.672 = 11.51632$$

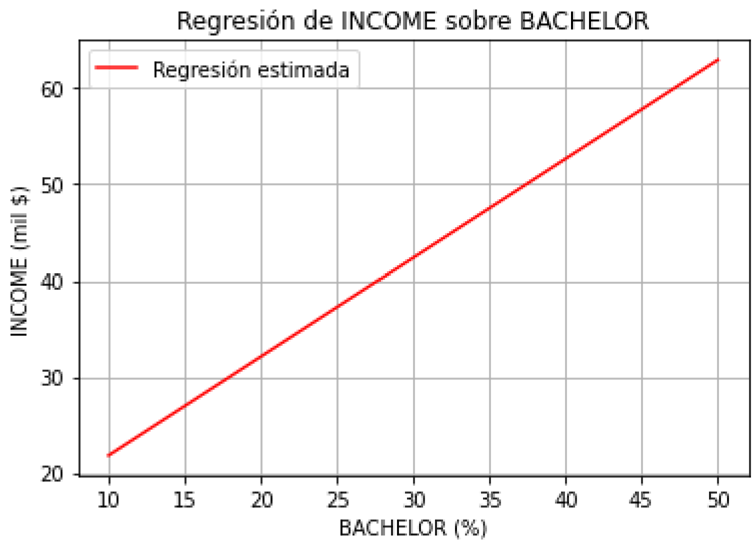
El intercepto estimado es $a \approx 11.52$.

---

b. Relación Estimada y Gráfica

La relación es **positiva y creciente** a una **tasa constante**. La ecuación de regresión estimada es:

$$\widehat{INCOME} = 11.52 + 1.029 \cdot BACHELOR$$



c) Error estándar de la pendiente

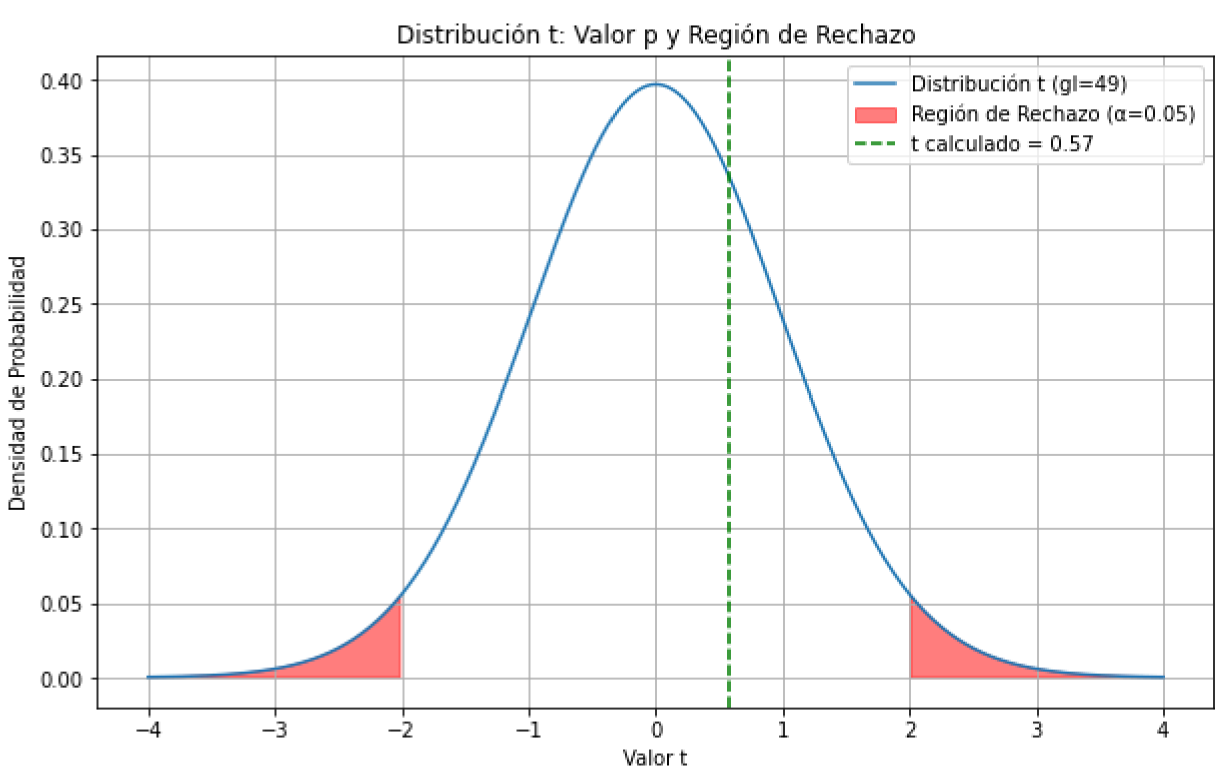
$$se(\hat{\beta}_1) = \frac{\hat{\beta}_1}{t_{BACHELOR}} = \frac{1.029}{10.75} \approx 0.096$$

d) t-estadístico para $H_0 : a = 10$

$$t = \frac{\hat{a} - 10}{se(a)} = \frac{11.51 - 10}{2.672} \approx 0.573$$

e) p-valor

$$p \approx 0.57 > 0.05 \Rightarrow \text{no rechazamos } H_0$$



f) Intervalo de confianza 99% para β_1

$$IC_{99\%} = 1.029 \pm 2.68 \cdot 0.096 \approx (0.772, 1.286)$$

Con un 95% de confianza los limites (0.772, 1.286) contienen al verdadero valor de β_1

g) Prueba de hipótesis para la pendiente

Se plantean las siguientes hipótesis con α=0.05:

$$H_0 : \beta_1 = 1$$
$$H_1 : \beta_1 \neq 1$$

Calculamos el estadístico *t*:

$$t = \frac{1.029 - 1}{0.096} = \frac{0.029}{0.096} = 0.302$$

Comparamos el valor absoluto del estadístico *t* calculado con el valor *t* crítico para $\alpha = 0.05$ y $gl = 49$, que es $**2.01**$.

$$|\ 0.302\ | < 2.01$$

Como el valor de nuestro estadístico *t* calculado es menor que el valor crítico, **no rechazamos la hipótesis nula**.

No hay evidencia estadística suficiente para concluir que la relación entre el porcentaje de graduados universitarios y el ingreso per cápita sea diferente de 1. Económicamente, esto significa que **no podemos descartar la posibilidad de que un aumento**

de un punto porcentual en la tasa de titulados universitarios se asocie con un aumento de exactamente $1,000 en el ingreso per cápita.

EJERCICIO-38

3.8 Using 2011 data on 141 U.S. public research universities, we examine the relationship between cost per student and full-time university enrollment. Let ACA = real academic cost per student (thousands of dollars), and let $FTESTU$ = full-time student enrollment (thousands of students). The least squares fitted relation is $\widehat{ACA} = 14.656 + 0.266FTESTU$.

a. For the regression, the 95% interval estimate for the intercept is [10.602, 18.710]. Calculate the standard error of the estimated intercept.

b. From the regression output, the standard error for the slope coefficient is 0.081. Test the null hypothesis that the true slope, β_2, is 0.25 (or less) against the alternative that the true slope is greater than 0.25 using the 10% level of significance. Show all steps of this hypothesis test, including the null and alternative hypotheses, and state your conclusion.

c. On the regression output, the automatically provided p-value for the estimated slope is 0.001. What is the meaning of this value? Use a sketch to illustrate your answer.

d. A member of the board of supervisors states that ACA should fall if we admit more students. Using the estimated equation and the information in parts (a)–(c), test the null hypothesis that the slope parameter β_2 is zero, or positive, against the alternative hypothesis that it is negative. Use the 5% level of significance. Show all steps of this hypothesis test, including the null and alternative hypotheses, and state your conclusion. Is there any statistical support for the board member's conjecture?

e. In 2011, Louisiana State University (LSU) had a full-time student enrollment of 27,950. Based on the estimated equation, the least squares estimate of $E(ACA|FTESTU = 27,950)$ is 22.079, with standard error 0.964. The actual value of ACA for LSU that year was 21.403. Would you say that this value is surprising or not surprising? Explain.

a. Calcular el error estándar del intercepto estimado

Se nos da el intervalo de confianza (IC) del 95% para el intercepto: $[10.602, 18.710]$.
La fórmula para el IC es:

$$IC = \hat{\beta}_1 \pm \underbrace{(t_{\text{crítico}} \times se(\hat{\beta}_1))}_{\text{Margen de error}}$$

El margen de error (ME) es la mitad del ancho del intervalo:

$$ME = \frac{\text{Límite Superior} - \text{Límite Inferior}}{2} = \frac{18.710 - 10.602}{2} = 4.054$$

Necesitamos el valor de $t_{\text{crítico}}$ para un 95% de confianza con $N - k - 1 = 141 - 1 - 1 = 139$ grados de libertad:

$$t_{\text{crítico}} \approx 1.977$$

Ahora, despejamos el error estándar:

$$se(\hat{\beta}_1) = \frac{ME}{t_{\text{crítico}}} = \frac{4.054}{1.977} \approx 2.051$$

Respuesta: $se(\hat{\beta}_1) \approx 2.051$.

---

b. Prueba de hipótesis para la pendiente $(\beta_2 > 0.25)$

Nivel de significancia: $\alpha = 0.10$

Hipótesis:

- $H_0 : \beta_2 \leq 0.25$
- $H_1 : \beta_2 > 0.25$

Estadístico de prueba:

$$t = \frac{\hat{\beta}_2 - \beta_{2,H0}}{se(\hat{\beta}_2)} = \frac{0.266 - 0.25}{0.081} \approx 0.198$$

Valor crítico: $t_{\text{crítico}} \approx 1.289$

Decisión: $0.198 < 1.289 \Rightarrow$ No rechazamos H_0.

Conclusión: No hay evidencia suficiente para afirmar que $\beta_2 > 0.25$.

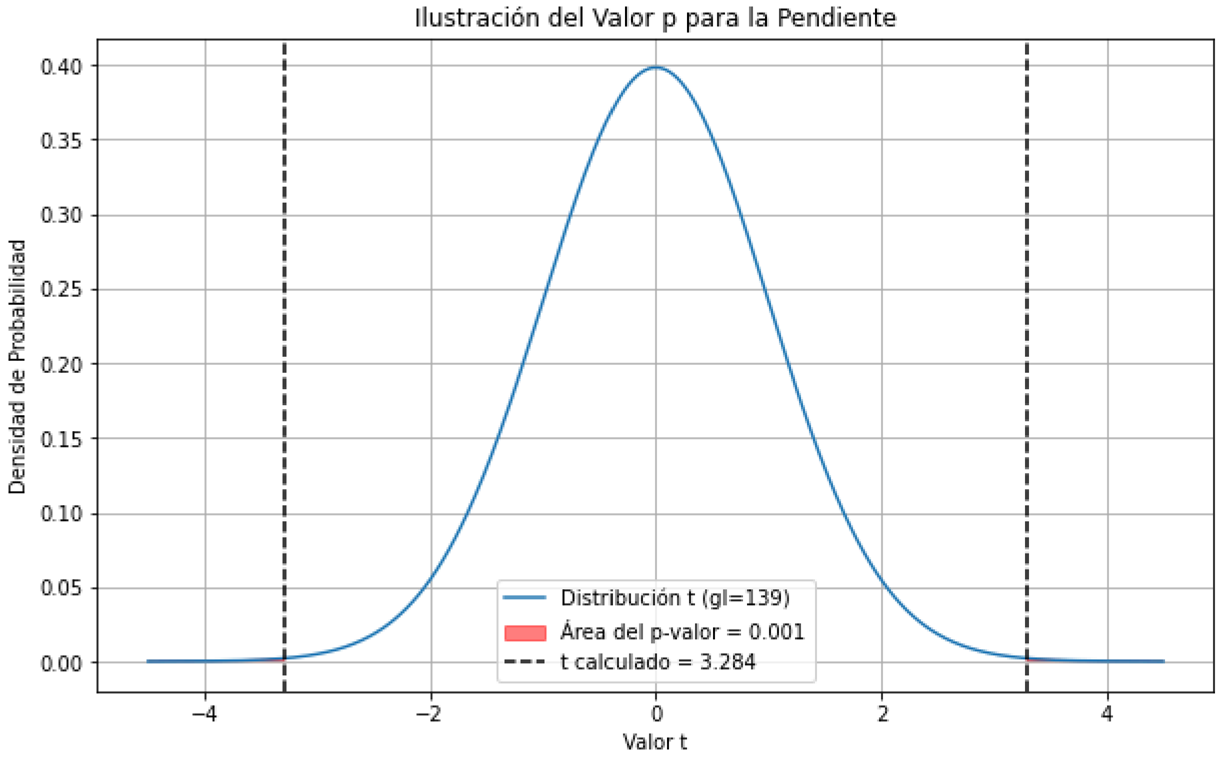
---

c. Significado del valor *p* de la pendiente

El valor $p = 0.001$ indica que, si $\beta_2 = 0$, la probabilidad de obtener un coeficiente tan extremo como 0.266 por azar es solo 0.1%. Como 0.001 es muy bajo, **rechazamos** H_0 y concluimos que existe una relación estadísticamente significativa entre matrícula y costo por estudiante.

Estadístico asociado:

$$t = \frac{0.266 - 0}{0.081} \approx 3.284$$



---

d. Prueba de la conjetura del miembro de la junta

Nivel de significancia: $\alpha = 0.05$, prueba de cola izquierda.

Hipótesis:

- $H_0 : \beta_2 \geq 0$
- $H_1 : \beta_2 < 0$

Estadístico de prueba:

$$t = \frac{0.266 - 0}{0.081} \approx 3.284$$

Valor crítico: $t_{\text{crítico}} \approx -1.656$

Decisión: $3.284 > -1.656 \Rightarrow$ No rechazamos H_0.

Conclusión: La pendiente positiva observada contradice la conjetura; no hay soporte estadístico para ella.

---

e. ¿Es sorprendente el valor de LSU?

Valor real: 21.403, valor predicho: 22.079, error estándar de predicción: 0.964.

Calculamos la "puntuación z":

$$\text{Puntuación} = \frac{21.403 - 22.079}{0.964} \approx -0.701$$

Interpretación: Como $|-0.701| < 2$, el valor de LSU **no es sorprendente**, rechazamos la hipótesis nula.

EJERCICIO-310

3.10 Using data from 2013 on 64 black females, the estimated log-linear regression between $WAGE$ (earnings per hour, in \$) and years of education, $EDUC$ is $\widehat{\ln(WAGE)} = 1.58 + 0.09EDUC$. The reported t-statistic for the slope coefficient is 3.95.

a. Test at the 5% level of significance, the null hypothesis that the return to an additional year of education is less than or equal to 8% against the alternative that the rate of return to education is more than 8%. In your answer, show (i) the formal null and alternative hypotheses, (ii) the test statistic and its distribution under the null hypothesis, (iii) the rejection region (in a figure), (iv) the calculated value of the test statistic, and (v) state your conclusion, with its economic interpretation.

b. Testing the null hypothesis that the return to education is 8%, against the alternative that it is not 8%, we obtain the p-value 0.684. What is the p-value for the test in part (a)? In a sketch, show for the test in part (a) the p-value and the 5% critical value from the t-distribution.

c. Construct a 90% interval estimate for the return to an additional year of education and state its interpretation.

a. Prueba de hipótesis sobre el retorno a la educación

Nivel de significancia: $\alpha = 0.05$

(i) Hipótesis Nula y Alternativa

- $H_0 : \beta_1 \leq 0.08$
- $H_1 : \beta_1 > 0.08$

(ii) Estadístico de Prueba y Grados de Libertad

$$gl = N - k - 1 = 64 - 1 - 1 = 62$$

(iii) Región de Rechazo

Para $\alpha = 0.05$, prueba de cola derecha:

$$t_{\text{crítico}} \approx 1.670$$

Rechazamos H_0 si $t_{\text{calculado}} > 1.670$.

(iv) Cálculo del Estadístico de Prueba

Primero, obtenemos el error estándar de la pendiente:

$$se(\hat{\beta}_1) = \frac{\hat{\beta}_1}{t_{\text{reportado}}} = \frac{0.09}{3.95} \approx 0.02278$$

Ahora calculamos el estadístico *t* para $H_0 : \beta_1 = 0.08$:

$$t_{\text{calculado}} = \frac{\hat{\beta}_1 - \beta_{1,H0}}{se(\hat{\beta}_1)} = \frac{0.09 - 0.08}{0.02278} \approx 0.439$$

(v) Conclusión e Interpretación

Como $0.439 < 1.670$, **no rechazamos** H_0.
Interpretación: No hay evidencia suficiente para afirmar que el retorno a la educación sea mayor al 8% para las mujeres negras.

---

b. Valor *p* para la prueba

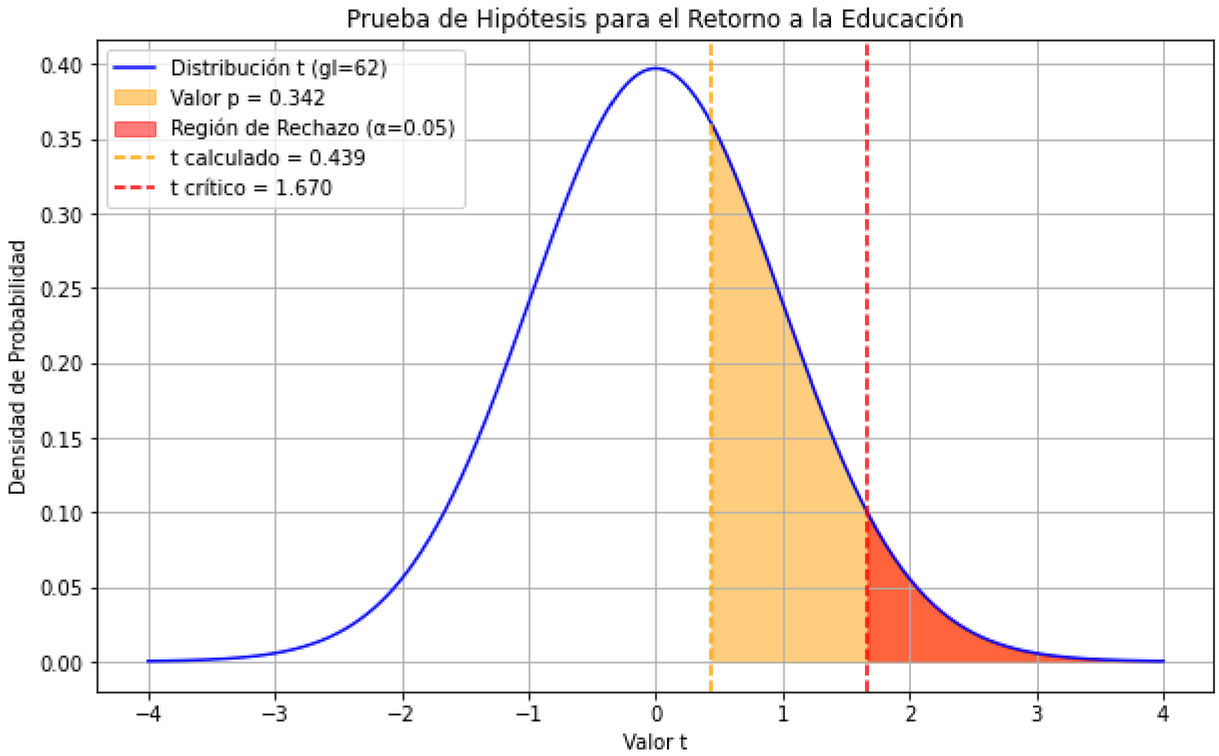
Se da un valor *p* de 0.684 para una prueba de **dos colas**. Para nuestra prueba de **cola derecha**:

$$p_{\text{una cola}} = \frac{0.684}{2} = 0.342$$

Como $0.342 > 0.05$, esto confirma que **no rechazamos** H_0.

Bosquejo:

- Marcamos $t_{\text{calculado}} = 0.439$; el área a su derecha es $p = 0.342$.
- Marcamos $t_{\text{crítico}} = 1.670$; el área a su derecha es la región de rechazo.
- Se observa que $t_{\text{calculado}}$ no cae en la región de rechazo.



---

c. Intervalo de Confianza del 90%

Fórmula general:

$$IC = \hat{\beta}_1 \pm (t_{\text{crítico}} \cdot se(\hat{\beta}_1))$$

Para un 90% de confianza y $gl = 62$:

$$t_{\text{crítico}} \approx 1.670$$

$$\text{Margen de Error} = 1.670 \times 0.02278 \approx 0.03807$$

Límites del IC:

$$\text{Límite Inferior} = 0.09 - 0.03807 \approx 0.0519$$

$$\text{Límite Superior} = 0.09 + 0.03807 \approx 0.1281$$

Intervalo de Confianza del 90%:

$$[0.052, 0.128]$$

Interpretación: Con un 90% de confianza, el retorno a un año adicional de educación para mujeres negras se encuentra entre un **5.2% y 12.8%** de incremento en los ingresos por hora.

EJERCICIO-321

3.21 The capital asset pricing model (CAPM) is described in Exercise 2.16. Use all available observations in the data file *capm5* for this exercise.

a. Construct 95% interval estimates of Exxon-Mobil's and Microsoft's "beta." Assume that you are a stockbroker. Explain these results to an investor who has come to you for advice.

b. Test at the 5% level of significance the hypothesis that Ford's "beta" value is one against the alternative that it is not equal to one. What is the economic interpretation of a beta equal to one? Repeat the test and state your conclusions for General Electric's stock and Exxon-Mobil's stock. Clearly state the test statistic used and the rejection region for each test, and compute the *p*-value.

c. Test at the 5% level of significance the null hypothesis that Exxon-Mobil's "beta" value is greater than or equal to one against the alternative that it is less than one. Clearly state the test statistic used and the rejection region for each test, and compute the *p*-value. What is the economic interpretation of a beta less than one?

d. Test at the 5% level of significance the null hypothesis that Microsoft's "beta" value is less than or equal to one against the alternative that it is greater than one. Clearly state the test statistic used and the rejection region for each test, and compute the *p*-value. What is the economic interpretation of a beta more than one?

e. Test at the 5% significance level, the null hypothesis that the intercept term in the CAPM model for Ford's stock is zero, against the alternative that it is not. What do you conclude? Repeat the test and state your conclusions for General Electric's stock and Exxon-Mobil's stock. Clearly state the test statistic used and the rejection region for each test, and compute the *p*-value.

a. Intervalos de Confianza para Beta (Exxon-Mobil y Microsoft)

El **beta** (β) de una acción mide su volatilidad en relación con el mercado. Un beta > 1 significa que la acción es más volátil que el mercado; un beta < 1 significa que es menos volátil.

Explicación para un inversionista:

- **Exxon-Mobil:** "Tenemos una confianza del 95% de que el beta de Exxon-Mobil se encuentra entre **0.315 y 0.598**. Como todo este rango está muy por debajo de 1, consideramos a Exxon-Mobil una **acción defensiva**. Esto significa que tiende a moverse con menos intensidad que el mercado general, lo cual puede ser atractivo en tiempos de incertidumbre".
- **Microsoft:** "Para Microsoft, el intervalo de confianza del 95% para su beta es de **[0.961, 1.443]**. Este rango incluye el valor de 1, pero la mayor parte del intervalo está por encima. Esto sugiere que Microsoft es una **acción moderadamente agresiva**, con una volatilidad probablemente un poco mayor a la del mercado en general".

---

b. Prueba de Hipótesis: Beta = 1

Interpretación económica: Un beta igual a 1 significa que una acción tiene, en promedio, el mismo riesgo sistemático que el mercado. Se espera que se mueva en la misma dirección y con la misma magnitud que el índice de mercado.

- **Ford:** El valor *p* de la prueba es **0.0016**. Como $0.0016 < 0.05$, **rechazamos la hipótesis nula**. El beta de Ford es estadísticamente diferente de 1.
- **General Electric:** El valor *p* de la prueba es **0.1002**. Como $0.1002 > 0.05$, **no rechazamos la hipótesis nula**. No tenemos evidencia suficiente para decir que el beta de GE sea diferente de 1.
- **Exxon-Mobil:** El valor *p* de la prueba es **0.0000**. Como $0.0000 < 0.05$, **rechazamos la hipótesis nula**. El beta de Exxon-Mobil es estadísticamente diferente de 1.

---

c. Prueba de Hipótesis: Beta de Exxon-Mobil < 1

- **Hipótesis:** $H_0 : \beta_{XOM} \geq 1 \quad \text{vs.} \quad H_1 : \beta_{XOM} < 1$
- **Estadístico de prueba:** $t = \frac{0.4565-1}{0.07155} \approx -7.60$
- **Región de rechazo:** Para una prueba de cola izquierda con $\alpha = 0.05$ y $gl = 178$, el valor crítico es $t_{\text{crítico}} \approx -1.653$. Rechazamos H_0 si $t < -1.653$.
- **Valor-p:** El valor *p* de una cola es $p < 0.0005$.
- **Conclusión:** Dado que $-7.60 < -1.653$, **rechazamos la hipótesis nula**. Hay evidencia muy fuerte de que el beta de Exxon-Mobil es menor que 1, confirmando su perfil como **acción defensiva**.

---

d. Prueba de Hipótesis: Beta de Microsoft > 1

- **Hipótesis:** $H_0 : \beta_{MSFT} \leq 1 \quad \text{vs.} \quad H_1 : \beta_{MSFT} > 1$
- **Estadístico de prueba:** $t = \frac{1.2018-1}{0.12215} \approx 1.652$
- **Región de rechazo:** Para una prueba de cola derecha con $\alpha = 0.05$ y $gl = 178$, el valor crítico es $t_{\text{crítico}} \approx 1.653$. Rechazamos H_0 si $t > 1.653$.
- **Valor-p:** $P(T_{178} > 1.652) \approx 0.0504$.
- **Conclusión:** Este es un caso límite. Técnicamente, como $1.652 < 1.653$ y el valor *p* es ligeramente mayor a 0.05, **no rechazamos la hipótesis nula**. No tenemos la evidencia suficiente para afirmar, con un 5% de significancia, que el beta de Microsoft sea mayor que 1.
- **Interpretación económica:** Un beta mayor a 1 caracteriza a una **acción agresiva**, que tiende a tener mayores ganancias que el mercado en épocas de bonanza, pero también mayores pérdidas en las caídas.

---

e. Prueba de Hipótesis: Intercepto (Alfa) = 0

El **alfa** (α) mide el "rendimiento anormal" de una acción que no es explicado por el mercado. En un mercado eficiente, se espera que el alfa sea cero.

- **Ford:** El estadístico *t* para el intercepto es **0.37** y el valor *p* es **0.712**. Como $p > 0.05$, **no rechazamos la hipótesis nula**.
- **General Electric:** El estadístico *t* es **-0.22** y el valor *p* es **0.829**. Como $p > 0.05$, **no rechazamos la hipótesis nula**.
- **Exxon-Mobil:** El estadístico *t* es **1.49** y el valor *p* es **0.137**. Como $p > 0.05$, **no rechazamos la hipótesis nula**.

Conclusión general: Para las tres acciones, no encontramos evidencia de que hayan generado rendimientos anormales (positivos o negativos) después de ajustar por el riesgo del mercado. Esto es consistente con la teoría del CAPM.

EJERCICIO-322

3.22 The data file *collegetown* contains data on 500 single-family houses sold in Baton Rouge, Louisiana, during 2009–2013. The data include sale price (in \$1000 units), *PRICE*, and total interior area in hundreds of square feet, *SQFT*.

a. Using the linear regression $PRICE = \beta_1 + \beta_2 SQFT + e$, estimate the elasticity of expected house *PRICE* with respect to *SQFT*, evaluated at the sample means. Construct a 95% interval estimate for the elasticity, treating the sample means as if they are given (not random) numbers. What is the interpretation of the interval?

b. Test the null hypothesis that the elasticity, calculated in part (a), is one against the alternative that the elasticity is not one. Use the 1% level of significance. Clearly state the test statistic used, the rejection region, and the test p-value. What do you conclude?

c. Using the linear regression model $PRICE = \beta_1 + \beta_2 SQFT + e$, test the hypothesis that the marginal effect on expected house price of increasing house size by 100 square feet is less than or equal to \$13000 against the alternative that the marginal effect will be greater than \$13000. Use the 5% level of significance. Clearly state the test statistic used, the rejection region, and the test p-value. What do you conclude?

d. Using the linear regression $PRICE = \beta_1 + \beta_2 SQFT + e$, estimate the expected price, $E(PRICE|SQFT) = \beta_1 + \beta_2 SOFT$, for a house of 2000 square feet. Construct a 95% interval estimate of the expected price. Describe your interval estimate to a general audience.

e. Locate houses in the sample with 2000 square feet of living area. Calculate the sample mean (average) of their selling prices. Is the sample average of the selling price for houses with $SQFT = 20$ compatible with the result in part (d)? Explain.

a. Elasticidad del precio y su intervalo de confianza

La elasticidad se calcula como:

$$\eta = \hat{\beta}_2 \times \frac{\overline{SQFT}}{\overline{PRICE}}$$

Con los datos:

$$\eta = 13.40294 \times \frac{27.28212}{250.2369} \approx 1.461$$

Resultados:

- Elasticidad estimada: **1.461**
- IC 95%: **[1.351, 1.571]**

Interpretación: Para una casa promedio, un aumento del 1% en los pies cuadrados se asocia con un aumento del 1.35% a 1.57% en el precio. Como el IC > 1, la relación es **elástica**.

---

b. Prueba de hipótesis para la elasticidad

Hipótesis:

- $H_0 : \eta = 1$
- $H_1 : \eta \neq 1$

Estadístico:

$$t = \frac{\hat{\eta} - \eta_{H0}}{se(\hat{\eta})} = \frac{1.461 - 1}{0.0558463} \approx 8.259$$

Valores críticos ($\alpha = 0.01$, $gl = 498$): $t_{\text{critico}} \approx \pm 2.586$.
Como $|8.259| > 2.586$ y $p < 0.001$, **rechazamos** H_0.

Conclusión: La elasticidad es estadísticamente distinta de 1.

---

c. Prueba de hipótesis para el efecto marginal

Queremos probar si el efecto de 100 pies² es mayor que 13 (miles de dólares).

Hipótesis:

- $H_0 : \beta_2 \leq 13$
- $H_1 : \beta_2 > 13$

Estadístico:

$$t = \frac{13.40294 - 13}{0.4491636} \approx 0.897$$

Valor crítico ($\alpha = 0.05$, $gl = 498$): $t_{\text{crítico}} \approx 1.648$.
Como $0.897 < 1.648$ y $p \approx 0.185 > 0.05$, **no rechazamos** H_0.

Conclusión: No hay evidencia suficiente para afirmar que el efecto marginal supere los $13,000.

---

d. Predicción de precio para una casa de 2000 pies²

Para $SQFT = 20$:

- Precio esperado: $152.6351 \Rightarrow \$152{,}635.10$
- IC 95%: $[141.5496, 163.7206] \Rightarrow [\$141{,}550,\ \$163{,}721]$

Interpretación: El precio promedio estimado para una casa de 2000 pies² es de aproximadamente **$152,600**, con un rango plausible de **$141,550 a $163,721**.

---

e. Comparación con casas de la muestra

En la muestra hay 3 casas con 2000 pies²:

- Precio promedio observado: $163.3333 \Rightarrow \$163{,}333.33$

Conclusión: Este valor cae dentro del IC obtenido en (d), por lo que es **compatible** con la predicción del modelo.

EJERCICIO-323

3.23 The data file *collegetown* contains data on 500 single-family houses sold in Baton Rouge, Louisiana, during 2009–2013. The data include sale price in $1000 units, *PRICE*, and total interior area in hundreds of square feet, *SQFT*.

a. Using the quadratic regression model, $PRICE = \alpha_1 + \alpha_2 SOFT^2 + e$, test the hypothesis that the marginal effect on expected house price of increasing the size of a 2000 square foot house by 100 square feet is less than or equal to $13000 against the alternative that the marginal effect will be greater than $13000. Use the 5% level of significance. Clearly state the test statistic used, the rejection region, and the test *p*-value. What do you conclude?

b. Using the quadratic regression model in part (a), test the hypothesis that the marginal effect on expected house price of increasing the size of a 4000 square foot house by 100 square feet is less than or equal to $13000 against the alternative that the marginal effect will be greater than $13000.

Use the 5% level of significance. Clearly state the test statistic used, the rejection region, and the test *p*-value. What do you conclude?

c. Using the quadratic regression model in part (a), estimate the expected price $E(PRICE|SQFT) = \alpha_1 + \alpha_2 SQFT^2$ for a house of 2000 square feet. Construct a 95% interval estimate of the expected price. Describe your interval estimate to a general audience.

d. Locate houses in the sample with 2000 square feet of living area. Calculate the sample mean (average) of their selling prices. Is the sample average of the selling price for houses with $SQFT = 20$ compatible with the result in part (c)? Explain.

a. Efecto marginal para una casa de 2000 pies²

Modelo cuadrático:

$$PRICE = \alpha_1 + \alpha_2 \cdot SQFT^2$$

Efecto marginal:

$$ME = \frac{d(PRICE)}{d(SQFT)} = 2 \cdot \alpha_2 \cdot SQFT$$

Para $SQFT = 20$:

$$ME = 40 \cdot \alpha_2$$

Hipótesis:

- $H_0 : (40 \cdot \alpha_2) \leq 13$
- $H_1 : (40 \cdot \alpha_2) > 13$

Resultados (`lincom`):

- Estimado = 7.38076
- Error estándar = 0.210232

Estadístico:

$$t = \frac{7.38076 - 13}{0.210232} \approx -26.728$$

Valor crítico ($\alpha = 0.05$, $gl = 498$): $t_{\text{crítico}} \approx 1.648$.

Conclusión: Como $-26.728 < 1.648$ y $p \approx 1$, **no rechazamos** H_0.
El efecto marginal es mucho menor a $13,000.

---

b. Efecto marginal para una casa de 4000 pies²

Para $SQFT = 40$:

$$ME = 80 \cdot \alpha_2$$

Hipótesis:

- $H_0 : (80 \cdot \alpha_2) \leq 13$
- $H_1 : (80 \cdot \alpha_2) > 13$

Resultados (`lincom`):

- Estimado = 14.76152
- Error estándar = 0.4204641

Estadístico:

$$t = \frac{14.76152 - 13}{0.4204641} \approx 4.189$$

Valor crítico: $t_{\text{crítico}} \approx 1.648$.

Conclusión: Como $4.189 > 1.648$ y $p < 0.001$, **rechazamos** H_0.
El efecto marginal supera significativamente los $13,000.

---

c. Predicción de precio para una casa de 2000 pies²

Para $SQFT = 20 \Rightarrow SQFT^2 = 400$:

Resultados (`margins`):

- Precio esperado = 167.3735 $\Rightarrow$ **$167,373.50**
- IC 95% = [158.0481, 176.6988] $\Rightarrow$ **[$158,048 ; $176,699]**

Interpretación: El precio promedio estimado es **$167,400**. Con 95% de confianza, el verdadero precio promedio está entre **$158,000 y $176,700**.

---

d. Comparación con casas de la muestra

En la muestra:

- 3 casas con 2000 pies²
- Precio promedio observado = **$163,333.33**

Conclusión: Este valor está **dentro** del IC de (c). El promedio muestral es consistente con la predicción poblacional.

EJERCICIO-324

3.24 We introduced Professor Ray C. Fair's model for explaining and predicting U.S. presidential elections in Exercise 2.23. Fair's data, 26 observations for the election years from 1916 to 2016, are in the data file *fair5*. The dependent variable is $VOTE$ = percentage share of the popular vote won by the Democratic party. Define $GROWTH = INCUMB \times growth\ rate$, where growth rate is the annual rate of change in real per capita GDP in the first three quarters of the election year. If Democrats are the incumbent party, then $INCUMB = 1$; if the Republicans are the incumbent party then $INCUMB = -1$.

a. Estimate the linear regression, $VOTE = \beta_1 + \beta_2 GROWTH + e$, using data from 1916 to 2016. Construct a 95% interval estimate of the effect of economic growth on expected $VOTE$. How would you describe your finding to a general audience?

b. The expected $VOTE$ in favor of the Democratic candidate is $E(VOTE|GROWTH) = \beta_1 + \beta_2 GROWTH$. Estimate $E(VOTE|GROWTH = 4)$ and construct a 95% interval estimate and a 99% interval estimate. Assume a Democratic incumbent is a candidate for a second presidential term. Is achieving a 4% growth rate enough to ensure a victory? Explain.

c. Test the hypothesis that when $INCUMB = 1$ economic growth has either a zero or negative effect on expected $VOTE$ against the alternative that economic growth has a positive effect on expected $VOTE$. Use the 1% level of significance. Clearly state the test statistic used, the rejection region, and the test p-value. What do you conclude?

d. Define $INFLAT = INCUMB \times$ inflation rate, where the inflation rate is the growth in prices over the first 15 quarters of an administration. Using the data from 1916 to 2016, and the model $VOTE = \alpha_1 + \alpha_2 INFLAT + e$, test the hypothesis that inflation has no effect against the alternative that it does have an effect. Use the 1% level of significance. State the test statistic used, the rejection region, and the test p-value and state your conclusion.

a. Efecto del Crecimiento Económico en el Voto

Regresión: `regress vote growth`

- Intervalo de confianza 95%: **[0.628, 1.300]**

Interpretación: Con 95% de confianza, cada punto porcentual adicional de crecimiento económico aumenta el voto del partido en el poder entre **0.63 y 1.30 puntos porcentuales**.
➡️ Evidencia clara de que una economía fuerte favorece al incumbente.

---

b. Predicción del Voto con Crecimiento del 4%

Para `growth = 4` con un incumbente Demócrata:

- Voto esperado: **52.52%**
- IC 95%: **[50.46%, 54.58%]**
- IC 99%: **[49.72%, 55.32%]**

Explicación:

- A nivel 95%: incluso el peor escenario (50.46%) implica victoria.
- A nivel 99%: el límite inferior baja a 49.72%, por lo que existe una pequeña probabilidad de derrota ajustada.
  ➡️ El 4% de crecimiento implica una **posición muy fuerte, pero no invencible**.

---

c. Prueba de Hipótesis: ¿Efecto positivo del Crecimiento?

Hipótesis:

- $H_0 : \beta_2 \leq 0$ (sin efecto positivo)
- $H_1 : \beta_2 > 0$ (efecto positivo)

Resultados:

- Estadístico $t = 5.93$
- Valor crítico (cola derecha, $\alpha = 0.01$, $gl = 24$): $t_{critico} \approx 2.492$
- Valor-p: $p < 0.0005$

Conclusión: $5.93 > 2.492$ y $p \ll 0.01$
➡️ **Rechazamos** H_0. Existe evidencia abrumadora de que el crecimiento económico tiene un efecto positivo sobre el voto al incumbente.

---

d. Prueba de Hipótesis: ¿Efecto de la Inflación?

Hipótesis:

- $H_0 : \alpha_2 = 0$ (sin efecto)
- $H_1 : \alpha_2 \neq 0$ (efecto ≠ 0)

Resultados:

- Estadístico $t = 0.70$
- Valores críticos (dos colas, $\alpha = 0.01$, $gl = 24$): $t_{crítico} \approx \pm 2.797$
- Valor-p = 0.489

Conclusión: $|0.70| < 2.797$ y $p \gg 0.01$
➡️ **No rechazamos** H_0. No hay evidencia de que la inflación afecte significativamente los resultados electorales.

EJERCICIO-325

3.25 Using data on the "Ashcan School," we have an opportunity to study the market for art. What factors determine the value of a work of art? Use the data in the file *ashcan_small*. [Note: the file *ashcan* contains more variables.]

a. Define $YEARS\_OLD = DATE\_AUCTN - CREATION$, which is the age of the painting at the time of its sale. Use data on works that sold ($SOLD = 1$) to estimate the regression $\ln(RHAMMER) = \beta_1 + \beta_2 YEARS\_OLD + e$. Construct a 95% interval estimate for the percentage change in real hammer price given that a work of art is another year old at the time of sale. [*Hint*: Review the discussion of equation (2.28).] Explain the result to a potential art buyer.

b. Test the null hypothesis that each additional year of age increases the "hammer price" by 2%, against the two-sided alternative. Use the 5% level of significance.

c. The variable *DREC* is an indicator variable taking the value one if a sale occurred during a recession and is zero otherwise. Use data on works that sold ($SOLD = 1$) to estimate the regression model $\ln(RHAMMER) = \alpha_1 + \alpha_2 DREC + e$. Construct a 95% interval estimate of the percentage reduction in hammer price when selling in a recession. Explain your finding to a client who is considering selling during a recessionary period.

d. Test the conjecture that selling a work of art during a recession reduces the hammer price by 2% or less, against the alternative that the reduction in hammer price is greater than 2%. Use the 5% level of significance. Clearly state the test statistic used, the rejection region, and the test *p*-value. What is your conclusion?

a. Efecto de la Antigüedad en el Valor del Arte

Cambio porcentual en el precio por cada año adicional de antigüedad (`nlcom`):

- IC 95%: **[0.87%, 3.26%]**

Interpretación:
Cada año adicional de antigüedad aumenta el precio de venta de una pintura entre **0.87% y 3.26%** (95% de confianza).
➡️ La antigüedad es un factor de apreciación constante en el mercado del arte.

---

b. Test: ¿Aumento del 2% por Año?

Hipótesis:

- $H_0 : \beta_2 = 0.02$
- $H_1 : \beta_2 \neq 0.02$

Resultados:

- Valor *p* = **0.9426**

Conclusión: $p \gg 0.05$
➡️ **No rechazamos** H_0. Los datos son perfectamente consistentes con que el valor aumenta un 2% por año.

---

c. Efecto de una Recesión en el Valor del Arte

Estimación del impacto de vender en recesión:

- IC 95%: **[-168.75%, -39.64%]**

Interpretación:
Vender en recesión reduce el precio de venta entre **168% y 39%** (95% de confianza).
➡️ Evidencia de un impacto muy negativo; mejor esperar a condiciones económicas más favorables.

---

d. Test: ¿Reducción Mayor al 2%?

Hipótesis:

- $H_0 : \alpha_2 \geq -0.02$ (reducción ≤ 2%)
- $H_1 : \alpha_2 < -0.02$ (reducción > 2%)

Resultados:

- Estadístico $t = -3.112$
- Valor crítico ($\alpha = 0.05$, cola izq., $gl = 420$): $t_{crítico} \approx -1.649$
- Valor *p* ≈ 0.001

Conclusión: $-3.112 < -1.649$ y $p \ll 0.05$
➡️ **Rechazamos** H_0. La reducción en el precio en recesión es **significativamente mayor al 2%**.

EJERCICIO-326

3.26 How much does experience affect wage rates? The data file *cps5_small* contains 1200 observations on hourly wage rates, experience, and other variables from the March 2013 Current Population Survey (CPS). [Note: The data file *cps5* contains more observations and variables.]

a. Estimate the linear regression $WAGE = \beta_1 + \beta_2 EXPER + e$ and discuss the results.

b. Test the statistical significance of the estimated relationship at the 5% level. Use a one-tail test. What is your alternative hypothesis? What do you conclude?

c. Estimate the linear regression $WAGE = \beta_1 + \beta_2 EXPER + e$ for individuals living in a metropolitan area, where $METRO = 1$. Is there a statistically significant positive relationship between expected wages and experience at the 1% level? How much of an effect is there?

d. Estimate the linear regression $WAGE = \beta_1 + \beta_2 EXPER + e$ for individuals not living in a metropolitan area, where $METRO = 0$. Is there a statistically significant positive relationship between expected wages and experience at the 1% level? Can we safely say that experience has no effect on wages for individuals living in nonmetropolitan areas? Explain.

a. Regresión Lineal y Discusión de Resultados

Modelo estimado (toda la muestra):

$$\widehat{WAGE} = 21.62 + 0.086 \cdot EXPER$$

- **Coef. experiencia (0.086):** positivo y significativo ($p = 0.009$).
  ➡️ Cada año extra de experiencia aumenta el salario por hora en **8.6 centavos**.
- **Intercepto (21.62):** salario esperado con 0 años de experiencia.
- **Ajuste del modelo:** $R^2 = 0.0057$ → la experiencia explica solo ~0.6% de la variación en salarios.

---

b. Test de Significación Estadística

Hipótesis:

- $H_0 : \beta_2 \leq 0$
- $H_1 : \beta_2 > 0$

Resultados:

- Estadístico $t = 2.61$
- Valor *p* (dos colas) = 0.009
- Valor *p* (una cola) = 0.0045

Conclusión: $p < 0.05$
➡️ **Rechazamos** H_0. Existe evidencia sólida de que la experiencia tiene un efecto positivo sobre el salario.

---

c. Regresión para Áreas Metropolitanas

Modelo (986 observaciones):

$$\widehat{WAGE}_{\text{metro}} = 21.89 + 0.113 \cdot EXPER$$

- Coef. experiencia = 0.113
- Valor *p* (dos colas) = 0.003 → una cola = 0.0015 < 0.01

- **Conclusión:** Relación positiva y significativa al 1%.
  ➡️ Cada año extra de experiencia aumenta el salario por hora en **11.3 centavos** en áreas metropolitanas.

---

d. Regresión para Áreas No Metropolitanas

Modelo (214 observaciones):

$$\widehat{WAGE}_{\text{no-metro}} = 19.41 + 0.013 \cdot EXPER$$

- Coef. experiencia = 0.013
- Valor *p* (dos colas) = 0.801 → una cola = 0.4005 ≫ 0.01
- **Conclusión:** No hay evidencia de relación significativa.

Nota importante:

- Estimación = 1.3 centavos por año.
- IC 95% ≈ [-9¢, +11.6¢].
- El intervalo incluye 0 → no se puede rechazar H_0.
  ➡️ Lo correcto es decir que **no encontramos evidencia estadística suficiente** de un efecto, no que el efecto sea inexistente.

EJERCICIO-327

3.27 Is the relationship between experience and wages constant over one's lifetime? We will investigate this question using a quadratic model. The data file *cps5_small* contains 1200 observations on hourly wage rates, experience, and other variables from the March 2013 Current Population Survey (CPS). [Note: the data file *cps5* contains more observations and variables.]

a. Create the variable $EXPER30 = EXPER - 30$. Describe this variable. When is it positive, negative or zero?

b. Estimate by least squares the quadratic model $WAGE = \gamma_1 + \gamma_2(EXPER30)^2 + e$. Test the null hypothesis that $\gamma_2 = 0$ against the alternative $\gamma_2 \neq 0$ at the 1% level of significance. Is there a statistically significant quadratic relationship between expected *WAGE* and *EXPER30*?

c. Create a plot of the fitted value $\widehat{WAGE} = \hat{\gamma}_1 + \hat{\gamma}_2(EXPER30)^2$, on the *y*-axis, versus *EXPER30*, on the *x*-axis. Up to the value $EXPER30 = 0$ is the slope of the plot constant, or is it increasing, or decreasing? Up to the value $EXPER30 = 0$ is the function increasing at an increasing rate or increasing at a decreasing rate?

d. If $y = a + bx^2$ then $dy/dx = 2bx$. Using this result, calculate the estimated slope of the fitted function $\widehat{WAGE} = \hat{\gamma}_1 + \hat{\gamma}_2(EXPER30)^2$, when $EXPER = 0$, when $EXPER = 10$, and when $EXPER = 20$.

e. Calculate the *t*-statistic for the null hypothesis that the slope of the function is zero, $H_0: 2\gamma_2 EXPER30 = 0$, when $EXPER = 0$, when $EXPER = 10$, and when $EXPER = 20$.

a. Variable `exper30`

La variable `exper30` mide los años de experiencia en relación con el punto de referencia de 30 años:

- Negativa si el trabajador tiene menos de 30 años de experiencia.
- Igual a 0 si el trabajador tiene exactamente 30 años de experiencia.
- Positiva si el trabajador tiene más de 30 años de experiencia.

En la muestra: rango desde -30 (recién ingresado) hasta +32 (62 años de experiencia).

---

b. Modelo Cuadrático y Prueba de Hipótesis

Modelo estimado:

$$\widehat{WAGE} = 25.94 - 0.0104 \cdot (EXPER30)^2$$

Hipótesis:

- $H_0 : \gamma_2 = 0$ (no hay efecto cuadrático)
- $H_1 : \gamma_2 \neq 0$ (sí hay efecto cuadrático)

Resultados:

- p-valor para `exper30_sq` = 0.000

Conclusión: $p < 0.01$
➡️ **Rechazamos** H_0. Existe evidencia estadística fuerte de una relación cuadrática entre salario y experiencia.

c. Forma de la Relación y Pendiente

- El coeficiente cuadrático (-0.0104) es **negativo** → la curva es una parábola abierta hacia abajo.
- Implica un perfil salarial típico: salarios suben con la experiencia al inicio, alcanzan un máximo y luego decrecen.

📌 La pendiente **no es constante**:

- Para trabajadores jóvenes (`exper30 < 0`): salarios crecen, pero cada año adicional aporta menos que el anterior → crecimiento **a tasa decreciente**.
- Después del máximo (`exper30 > 0`): salarios tienden a caer.

---

d. Pendiente en Diferentes Puntos

La pendiente se obtiene derivando:

$$\text{Pendiente} = 2 \cdot \hat{\gamma}_2 \cdot EXPER30 = -0.0209 \cdot EXPER30$$

- **EXPER = 0 (exper30 = -30):**
  Pendiente = $-0.0209 \times (-30) = 0.627$
  ➡️ Cada año extra aumenta el salario en **62.7¢ por hora**.
- **EXPER = 10 (exper30 = -20):**
  Pendiente = $-0.0209 \times (-20) = 0.418$
  ➡️ Incremento de **41.8¢ por hora**.
- **EXPER = 20 (exper30 = -10):**
  Pendiente = $-0.0209 \times (-10) = 0.209$
  ➡️ Incremento de **20.9¢ por hora**.

---

e. Test de Hipótesis sobre la Pendiente

Hipótesis:

- $H_0 : 2 \cdot \gamma_2 \cdot EXPER30 = 0$
- $H_1 : 2 \cdot \gamma_2 \cdot EXPER30 \neq 0$

Como la pendiente depende de γ_2, probar la pendiente en cualquier punto es equivalente a probar $H_0 : \gamma_2 = 0$.

Resultado:

- Estadístico t de `exper30_sq` = **-5.81**
- Es el mismo en todos los puntos considerados.

Conclusión: fuerte evidencia de que la pendiente no es cero.

EJERCICIO-322

3.32 What is the relationship between crime and punishment? We use data from 90 North Carolina counties to examine the question. County crime rates and other characteristics are observed over the period 1981–1987. The data are in the file *crime*. Use the 1985 data for this exercise.

a. Calculate the summary statistics for *CRMRTE* (crimes committed per person) and *PRBARR* (the probability of arrest = the ratio of arrests to offenses), including the maximums and minimums. Does there appear to be much variation from county to county in these variables?

b. Plot *CRMRTE* versus *PRBARR*. Do you observe a relationship between these variables?

c. Estimate the linear regression model $CRMRTE = \beta_1 + \beta_2 PRBARR + e$. If we increase the probability of arrest by 10% what will be the effect on the crime rate? What is a 95% interval estimate of this quantity?

d. Test the null hypothesis that there is no relationship between the county crime rate and the probability of arrest versus the alternative that there is an inverse relationship. State the null and alternative hypotheses in terms of the model parameters. Clearly, state the test statistic and its distribution if the null hypothesis is true and the test rejection region. Use the 1% level of significance. What is your conclusion?

a. Estadísticas Descriptivas

Para 1985 en los 90 condados de Carolina del Norte:

- **Tasa de criminalidad (`crmrte`):**
  - Promedio: 0.0298 (29.8 crímenes por cada 1000 personas).

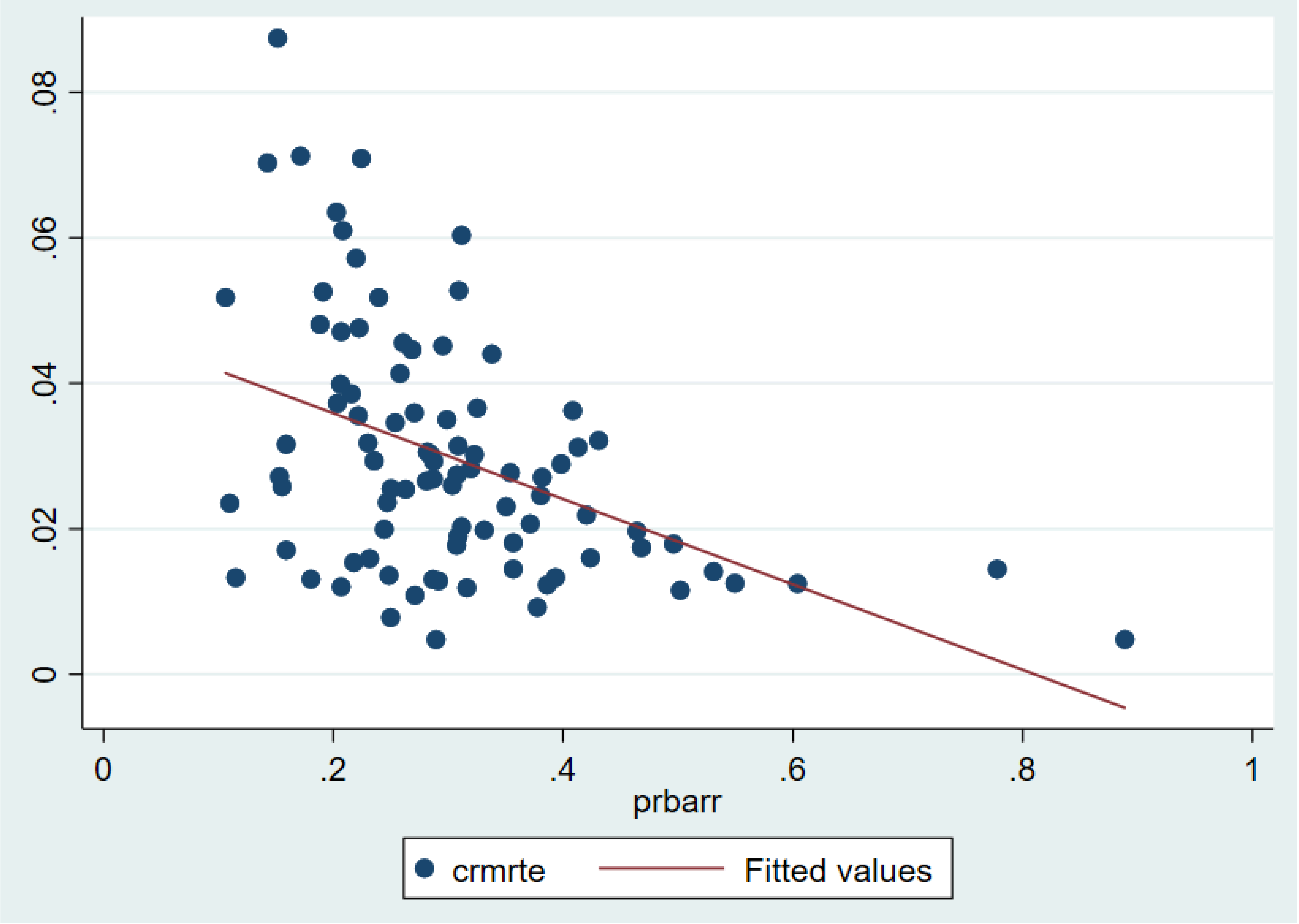
- Mínimo: 0.0048
- Máximo: 0.0875
- **Probabilidad de arresto (`prbarr`):**
  - Promedio: 0.3040 (30.4%).
  - Mínimo: 0.1062 (10.6%)
  - Máximo: 0.8889 (88.9%)

📌 **Variación:**

- El condado con más criminalidad tiene ~18 veces más crímenes per cápita que el de menor nivel.
- La probabilidad de arresto más alta es ~8 veces la más baja.
  ➡️ Hay gran disparidad entre condados.

---

b. Gráfica `crmrte` vs. `prbarr`



➡️ Los condados con mayor probabilidad de arresto tienden a tener **menos criminalidad**.

---

c. Estimación del Modelo Lineal

Modelo estimado:

$$\widehat{CRMRTE} = 0.0476 - 0.0588 \cdot PRBARR$$

- Aumento de 0.10 en `prbarr` → disminución de **0.00588** en `crmrte` .
- IC 95%: **[-0.00837, -0.00338]**

📌 **Interpretación:**
Si la probabilidad de arresto sube del 30% al 40%, la tasa de criminalidad baja en ~0.0059 delitos por persona. Con 95% de confianza, la reducción está entre 0.0034 y 0.0084.

---

d. Test de Hipótesis: Relación Inversa

Hipótesis:

- $H_0 : \beta_2 \geq 0$ (no hay relación inversa)
- $H_1 : \beta_2 < 0$ (sí hay relación inversa)

Resultados:

- Estadístico $t = -4.68$ (gl = 88)
- Valor crítico $(\alpha = 0.01)$ ≈ -2.37
- Valor *p* (dos colas) = 0.000 → una cola: $p < 0.0005$

Conclusión:
Como $-4.68 < -2.37$ y $p < 0.01$, **rechazamos** H_0.
➡️ Evidencia estadística muy fuerte de que **más probabilidad de arresto reduce la criminalidad**. 👮