Introduction to Econometrics
Chapter 4

Ezequiel Uriel Jiménez
University of Valencia

Valencia, September 2013

4 Hypothesis testing in the multiple regression model

4 Hypothesis testing in the multiple regression model

Motivation

Testing hypothesis can answer the following questions:

1. Is the marginal propensity to consume smaller than the average propensity to consume?
2. Has income a negative influence on infant mortality?
3. Does the rate of crime in an area plays a role in the prices of houses in that area?
4. Is the elasticity expenditure in fruit/income equal to 1? Is fruit a luxury good?
5. Is the Madrid stock exchange market efficient?
6. Is the rate of return of the Madrid Stock Exchange affected by the rate of return of the Tokyo Stock Exchange?
7. Are there constant returns to scale in the chemical industry?
8. Advertising or incentives?
9. Is the assumption of homogeneity admissible in the demand for fish?
10. Have tenure and age jointly a significant influence on wage?
11. Is the performance of a company crucial to set the salaries of CEOs?

All these questions are answered in this chapter

4.1 Hypothesis testing: an overview

TABLE 4.1. Some distributions used in hypothesis testing.

	1 restriction	*1 or more restrictions*
Known σ^2	*N*	*Chi-square*
Unknown σ^2	Student's *t*	Snedecor's *F*

4.1 Hypothesis testing: an overview

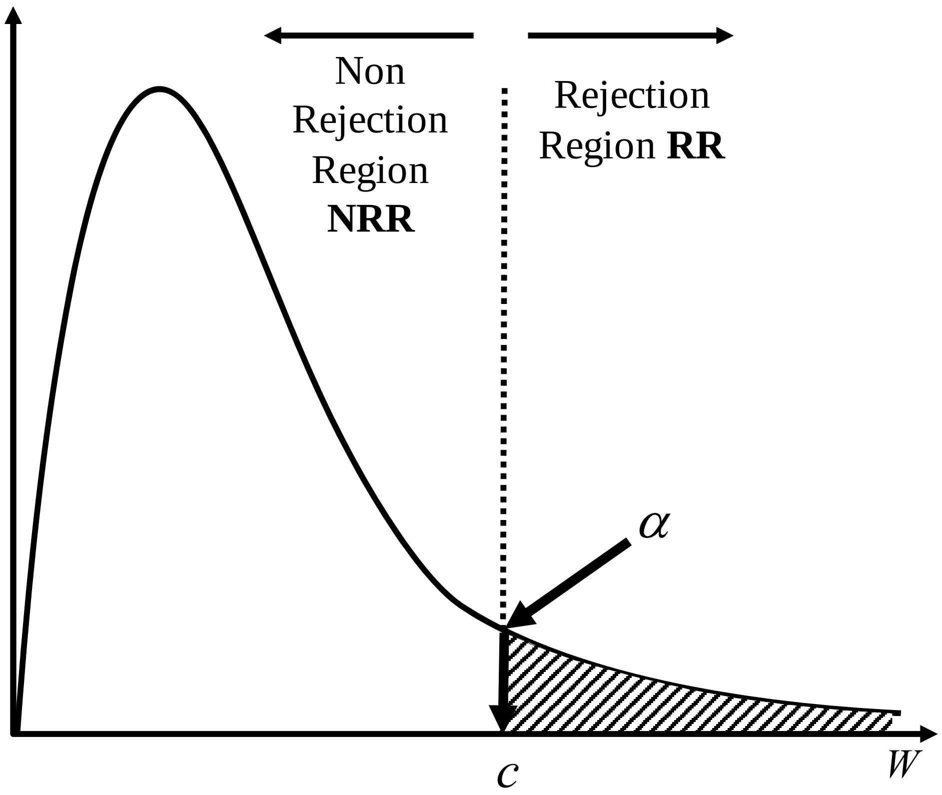

FIGURE 4.1. Hypothesis testing: classical approach.

4.2 Testing hypotheses using the *t* test

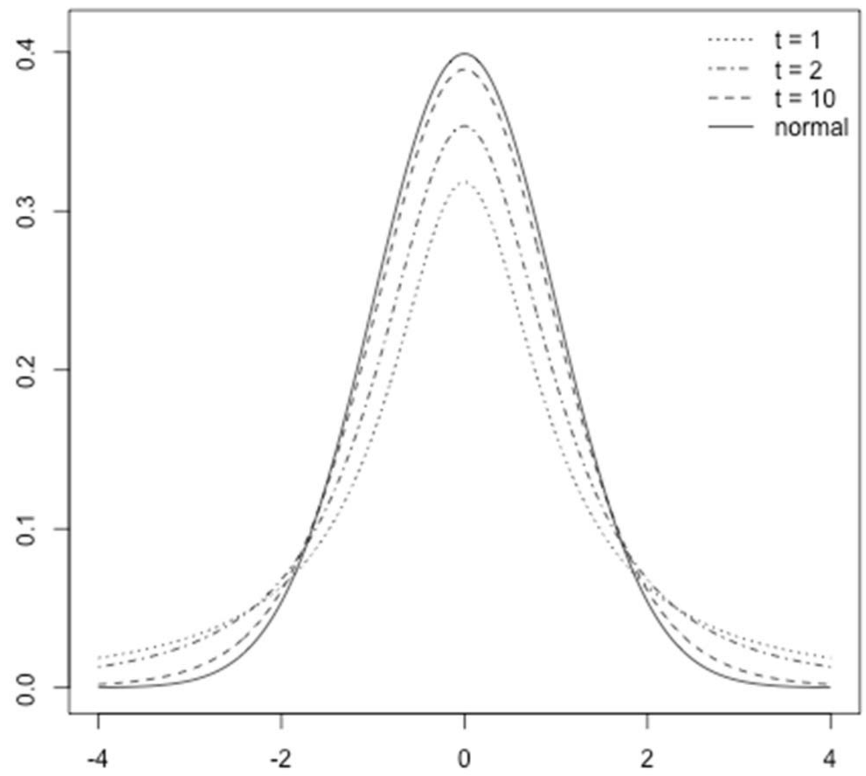

FIGURE 4. 2. Density functions: normal and *t* for different degrees of freedom.

4.2 Testing hypotheses using the *t* test

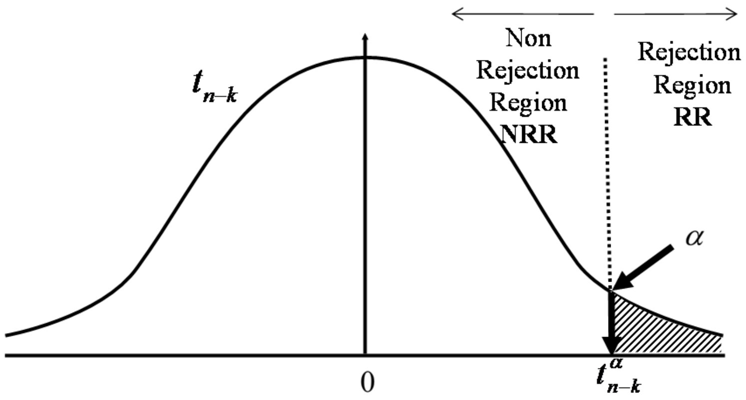

FIGURE 4.3. Rejection region using *t*: right-tail alternative hypothesis.

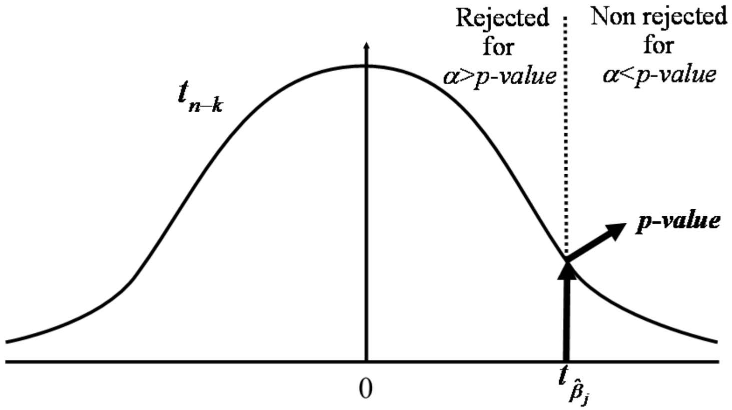

FIGURE 4.4. *p-value* using *t*: right-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.1 Is the marginal propensity to consume smaller than the average propensity to consume?

$$cons = \beta_1 + \beta_2 inc + u$$

$$\widehat{cons}_i = \underset{(0.350)}{0.41} + \underset{(0.062)}{0.843} inc_i$$

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 > 0$$

$$t = \frac{\hat{\beta}_1 - \beta_1^0}{se(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)} = \frac{0.41}{0.35} = 1.171$$

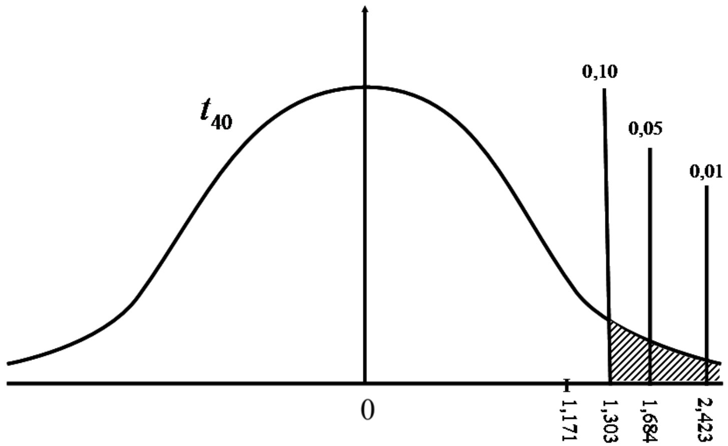

FIGURE 4.5. Example 4.1: Rejection region using *t* with a right-tail alternative hypothesis.

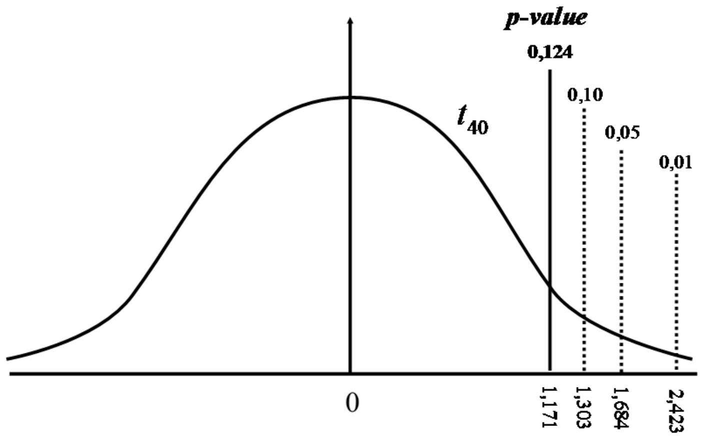

FIGURE 4.6. Example 4.1: *p-value* using *t* with right-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

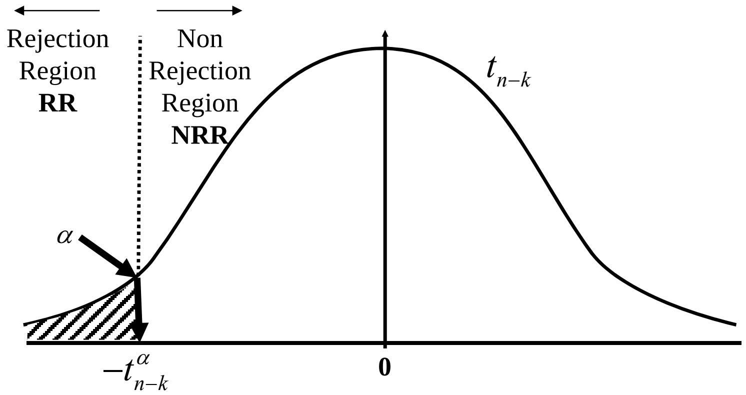

FIGURE 4.7. Rejection region using *t*: left-tail alternative hypothesis

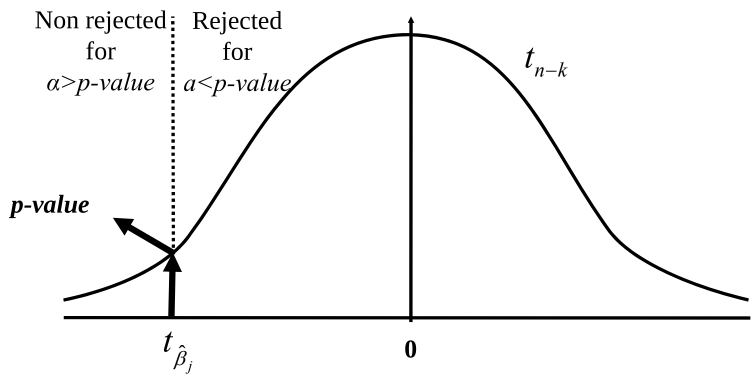

FIGURE 4.8. *p-value* using *t*: left-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.2 ***Has income a negative influence on infant mortality?***

$$deathun5 = \beta_1 + \beta_2 gnipc + \beta_3 ilitrate + u$$

$$\widehat{deathun5_i} = \underset{(5.93)}{27.91} - \underset{(0.00028)}{0.000826}\, gnipc_i + \underset{(0.183)}{2.043}\, ilitrate_i$$

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 < 0$$

$$t = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \frac{-0.000826}{0.00028} = -2.966$$

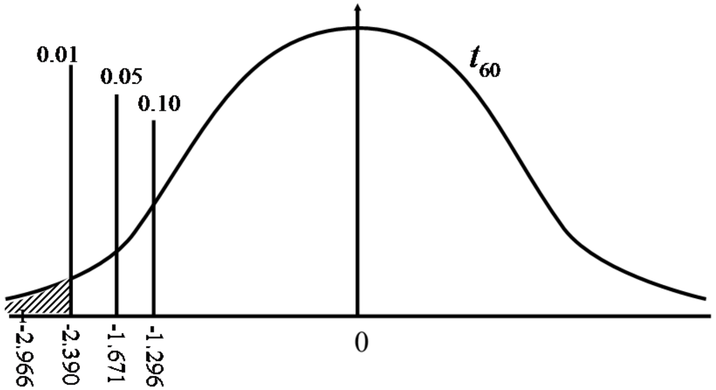

FIGURE 4.9. Example 4.2: Rejection region using *t* with a left-tail alternative hypothesis.

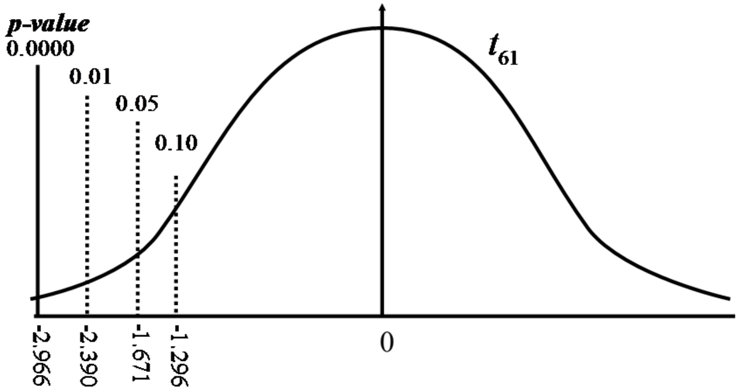

FIGURE 4.10. Example 4.2: *p-value* using *t* with a left-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

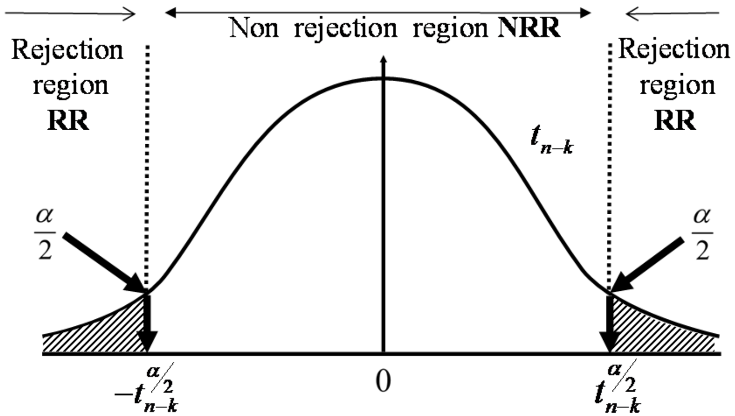

FIGURE 4.11. Rejection region using *t*: two-tail alternative hypothesis.

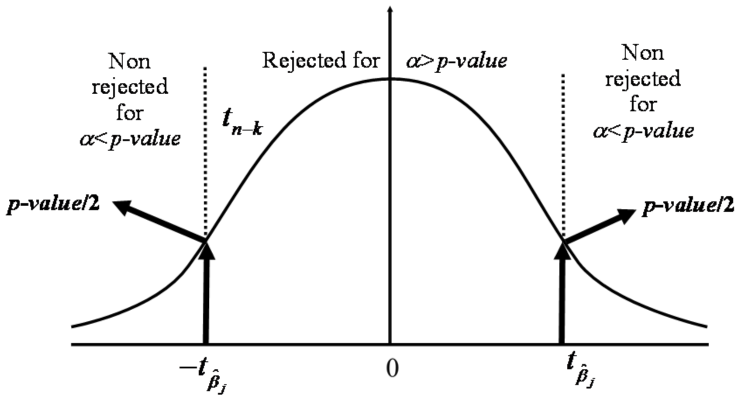

FIGURE 4.12. *p-value* using *t*: two-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.3 ***Has the rate of crime play a role in the price of houses in an area?***

$$price = \beta_1 + \beta_2 rooms + \beta_3 lowstat + \beta_4 crime + u$$

$$\widehat{price_i} = \underset{(8022)}{-15694} + \underset{(1210)}{6788}\, rooms_i - \underset{(80.7)}{268.2}\, lowstat_i - \underset{(960)}{3854}\, crime_i$$

TABLE 4.2. Standard output in the regression explaining house price. n=55.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
C	-15693.61	8021.989	-1.956324	0.0559
rooms	6788.401	1210.72	5.60691	0.0000
lowstat	-268.1636	80.70678	-3.32269	0.0017
crime	-3853.564	959.5618	-4.015962	0.0002

$$H_0 : \beta_4 = 0$$
$$H_1 : \beta_4 \neq 0$$

$$t = \frac{\hat{\beta}_4}{se(\hat{\beta}_4)} = \frac{-3854}{960} = -4.016$$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.3 ***Has the rate of crime play a role in the price of houses in an area?* (Continuation)**

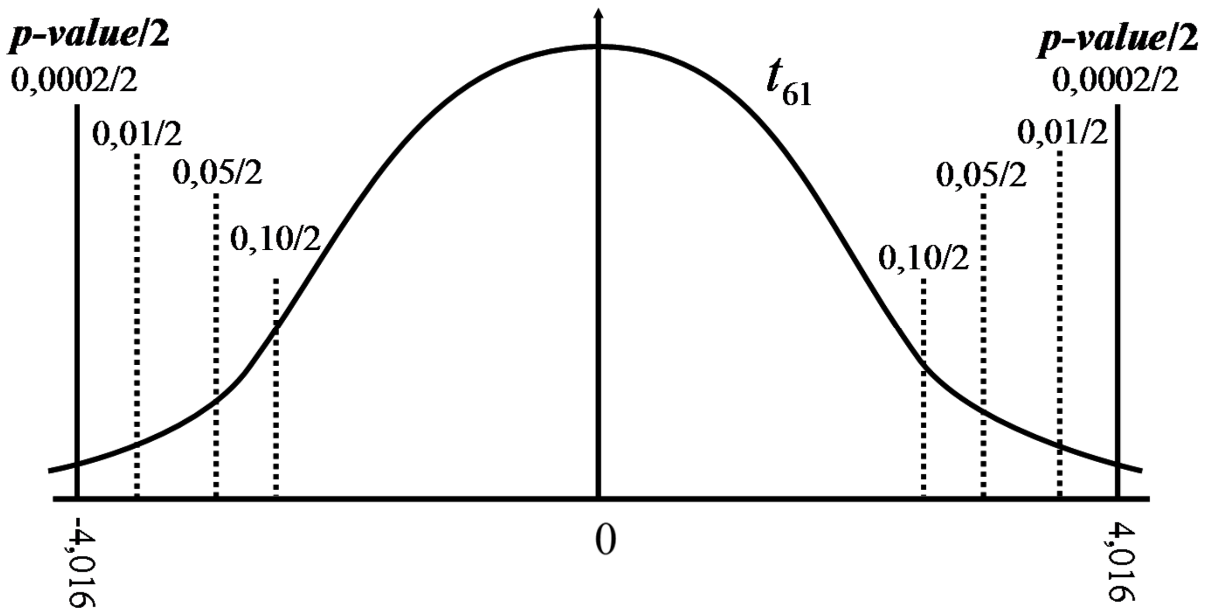

FIGURE 4.13. Example 4.3: *p-value* using *t* with a two-tail alternative hypothesis.

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.4 Is the elasticity expenditure in fruit/income equal to 1? Is fruit a luxury good?

$$\ln(fruit) = \beta_1 + \beta_2 \ln(inc) + \beta_3 househsize + \beta_4 punders + u$$

$$\widehat{\ln(fruit_i)} = \underset{(3.701)}{-9.768} + \underset{(0.512)}{2.005} \ln(inc_i) - \underset{(0.179)}{1.205}\, househsize_i - \underset{(0.013)}{0.018}\, punder5_i$$

TABLE 4.3. Standard output in a regression explaining expenditure in fruit. *n*=40.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
C	-9.767654	3.701469	-2.638859	0.0122
ln(inc)	2.004539	0.51237	3.912286	0.0004
househsize	-1.205348	0.178646	-6.747147	0.0000
punder5	-0.017946	0.013022	-1.378128	0.1767

$H_0 : \beta_2 = 1$

$H_1 : \beta_2 \neq 1$

$$t = \frac{\hat{\beta}_2 - \beta_2^0}{se(\hat{\beta}_2)} = \frac{\hat{\beta}_2 - 1}{se(\hat{\beta}_2)} = \frac{2.005 - 1}{0.512} = 1.961$$

$H_1 : \beta_2 > 1$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.5 Is the Madrid stock exchange market efficient?

Rate of total return: $RA_t = \dfrac{\Delta P_t + D_t + A_t}{P_{t-1}}$

Rate of return due to increase in quotation

Proportional change: $RA1_t = \dfrac{\Delta P_t}{P_{t-1}}$

Change in logarithms: $RA2_t = \Delta \ln P_t$

$$rmad92_t = \beta_1 + \beta_2 rmad92_{t-1} + u_t$$

$$\widehat{rmad92}_t = \underset{(0.0007)}{-0.0004} + \underset{(0.0629)}{0.1267}\, rmad92_{t-1}$$

$$R^2 = 0.0163 \quad n = 247$$

$$H_0 : \beta_2 = 1$$
$$H_1 : \beta_2 \neq 1$$

$$t = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \frac{0.1267}{0.0629} = 2.02$$

EXAMPLE 4.6 Is the rate of return of the Madrid Stock Exchange affected by the rate of return of the Tokyo Stock Exchange?

$$rmad92_t = \beta_1 + \beta_2 rtok92_t + u_t$$

$$\widehat{rmad92}_t = \underset{(0.0007)}{-0.0005} + \underset{(0.0375)}{0.1244}\, rtok92_t$$

$$R^2 = 0.0452 \quad n = 235$$

$$H_0 : \beta_2 = 1$$
$$H_1 : \beta_2 \neq 1$$

$$t = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \frac{0.1244}{0.0375} = 3.32$$

4.2 Testing hypotheses using the *t* test

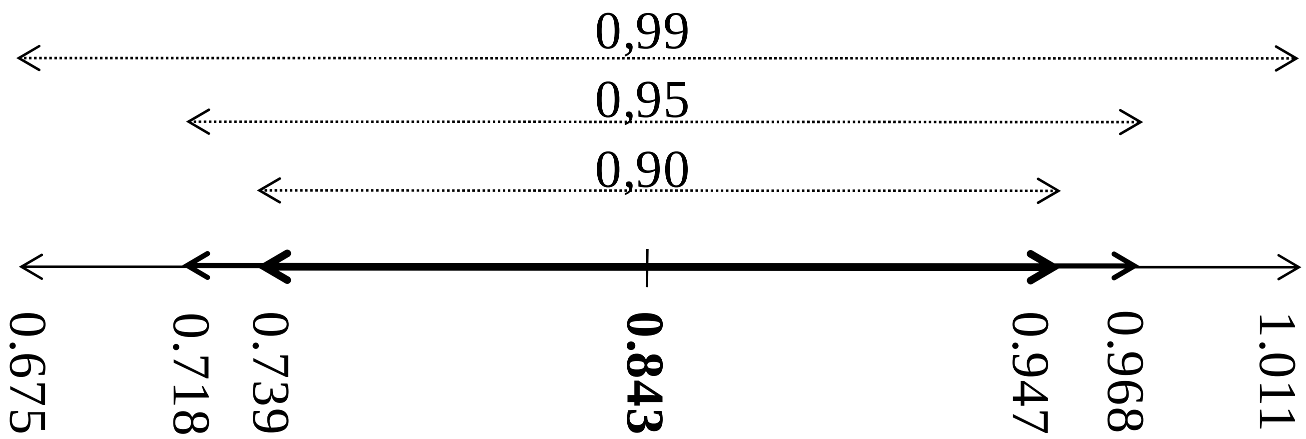

FIGURE 4.14. Confidence intervals for marginal propensity to consume in example 4.1.

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.7 Are there constant returns to scale in the chemical industry?

$$\ln(output) = \beta_1 + \beta_2 \ln(labor) + \beta_3 \ln(capital) + u$$

$$\widehat{\ln(output_i)} = \underset{(0.327)}{1.170} + \underset{(0.126)}{0.603} \ln(labor_i) + \underset{(0.085)}{0.376} \ln(capital_i)$$

TABLE 4.4. Standard output of the estimation of the production function: model (4-20).

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
constant	1.170644	0.326782	3.582339	0.0015
$\ln(labor)$	0.602999	0.125954	4.787457	0.0001
$\ln(capital)$	0.37571	0.085346	4.402204	0.0002

$$H_0: \beta_2 + \beta_3 = 1$$
$$H_1: \beta_2 + \beta_3 \neq 1$$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.7 Are there constant returns to scale in the chemical industry? (Cont.)

TABLE 4.5. Covariance matrix in the production function.

	constant	*ln(labor)*	*ln(capital)*
constant	0.106786	-0.019835	0.001189
ln(*labor*))	-0.019835	0.015864	-0.009616
ln(*capital*)	0.001189	-0.009616	0.007284

a) Procedure: using covariance matrix of estimators.

$$\widehat{\operatorname{var}(\hat{\beta}_2+\hat{\beta}_3)} = \widehat{\operatorname{var}(\hat{\beta}_2)} + \widehat{\operatorname{var}(\hat{\beta}_3)} + 2\times\widehat{\operatorname{covar}(\hat{\beta}_2,\hat{\beta}_3)}$$

$$se(\hat{\beta}_2+\hat{\beta}_3) = \sqrt{\widehat{\operatorname{var}(\hat{\beta}_2+\hat{\beta}_3)}} \qquad se(\hat{\beta}_2+\hat{\beta}_3) = \sqrt{0.015864+0.007284-2\times 0.009616} = 0.0626$$

$$t_{\hat{\beta}_2+\hat{\beta}_3} = \frac{\hat{\beta}_2+\hat{\beta}_3-1}{se(\hat{\beta}_2+\hat{\beta}_3)} = \frac{-0.02129}{0.0626} = -0.3402$$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.7 Are there constant returns to scale in the chemical industry? (Cont.)

b) Procedure: reparameterizing the model by introducing a new parameter .

$$\theta = \beta_2 + \beta_3 - 1 \Rightarrow \beta_2 = \theta - \beta_3 + 1$$

$$\ln(output) = \beta_1 + (\theta - \beta_3 + 1)\ln(labor) + \beta_3 \ln(capital) + u$$

$$\ln(output / labor) = \beta_1 + \theta \ln(labor) + \beta_3 \ln(capital / labor) + u$$

TABLE 4.6. Estimation output for the production function: reparameterized model.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
constant	1.170.644	0.326782	3.582.339	0.0015
ln(*labor*)	-0.021290	0.062577	-0.340227	0.7366
ln(*capital/labor*)	0.375710	0.085346	4402204	0.0002

$$H_0 : \theta_1 = 0$$

$$H_1 : \theta \neq 0$$

$$t = \frac{\hat{\theta}}{se(\hat{\theta})} = \frac{-0.02129}{0.0626} = -0.3402$$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.8 Advertising or incentives?

$$sales = \beta_1 + \beta_2 advert + \beta_3 incent + u$$

TABLE 4.7. Standard output of the regression for example 4.8.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
constant	396.5945	3548.111	0.111776	0.9125
advert	18.63673	8.924339	2.088304	0.0542
incent	30.69686	3.60442	8.516448	0.0000

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.8 Advertising or incentives? (Continuation)

TABLE 4.8. Covariance matrix for example 4.8.

	C	*advert*	*incent*
constant	12589095	-26674	-7101
advert	-26674	79.644	2.941
incent	-7101	2.941	12.992

$$H_0 : \beta_3 - \beta_2 = 0$$
$$H_1 : \beta_3 - \beta_2 > 0$$

$$se(\hat{\beta}_3 - \hat{\beta}_2) = \sqrt{79.644 + 12.992 - 2 \times 2.941} = 9.3142$$

$$t_{\hat{\beta}_3 - \hat{\beta}_2} = \frac{\hat{\beta}_3 - \hat{\beta}_2}{se(\hat{\beta}_3 - \hat{\beta}_2)} = \frac{30.697 - 18.637}{9.3142} = 1.295$$

4.2 Testing hypotheses using the *t* test

EXAMPLE 4.9 Testing the hypothesis of homogeneity in the demand for fish

$$\ln(fish) = \beta_1 + \beta_2 \ln(fishpr) + \beta_3 \ln(meatpr) + \beta_4 \ln(cons) + u$$

$$\widehat{\ln(fish_i)} = \underset{(2.30)}{7.788} - \underset{(0.133)}{0.460} \ln(fishpr_i) + \underset{(0.112)}{0.554} \ln(meatpr_i) + \underset{(0.137)}{0.322} \ln(cons_i)$$

Homogeneity resstriction:

$$\beta_2 + \beta_3 + \beta_4 = 0 \quad \Rightarrow \quad \theta = \beta_2 + \beta_3 + \beta_4$$

$$\ln(fish) = \beta_1 + \theta \ln(fishpr) + \beta_3 \ln(meatpr / fishpr) + \beta_4 \ln(cons / fishpr) + u$$

$$\widehat{\ln(fish_i)} = \underset{(2.30)}{7.788} - \underset{(0.1334)}{0.4596} \ln(fishpr_i) + \underset{(0.112)}{0.554} \ln(meatpr_i) + \underset{(0.137)}{0.322} \ln(cons_i)$$

$$t = \frac{\hat{\theta}}{se(\hat{\theta})} = \frac{-0.4596}{0.1334} = -3.44$$

4.3 Testing multiple linear restrictions using the *F* test

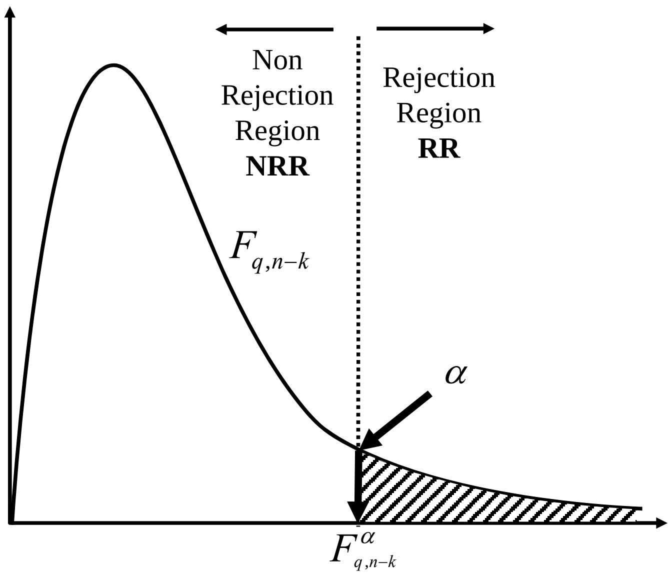

FIGURE 4.15. Rejection region and non rejection region using *F* distribution.

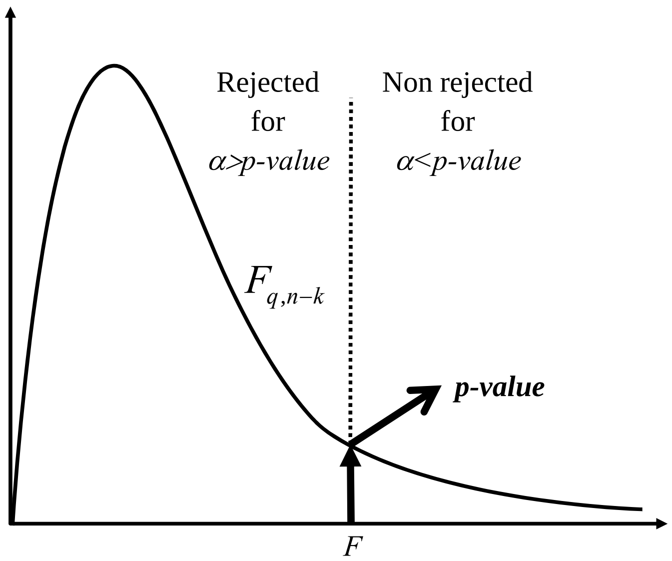

FIGURE 4.16. *p-value* using *F* distribution.

4.3 Testing multiple linear restrictions using the *F* test

EXAMPLE 4.10 Wage, experience, tenure and age

$$\ln(wage) = \beta_1 + \beta_2 educ + \beta_3 exper + \beta_4 tenure + \beta_5 age + u$$

$$\widehat{\ln(wage_i)} = 6.476 + 0.0658 educ_i + 0.0267 exper_i - 0.0094 tenure_i - 0.0209 age_i$$

$$RSS = 5.954$$

$$H_0 : \beta_4 = \beta_5 = 0$$

$$H_1 : H_0 \text{ is not true}$$

$$\ln(wage) = \beta_1 + \beta_2 educ + \beta_3 exper + u$$

$$\widehat{\ln(wage_i)} = 6.157 + 0.0457 educ_i + 0.0121 exper_i$$

$$RSS = 6.250$$

$$F = \frac{\left(RSS_R - RSS_{UR}\right)/q}{RSS_{UR}/(n-k)} = \frac{(6.250 - 5.954)/2}{5.954/48} = 1.193$$

4.3 Testing multiple linear restrictions using the *F* test

EXAMPLE 4.10 Wage, experience, tenure and age. (Continuation)

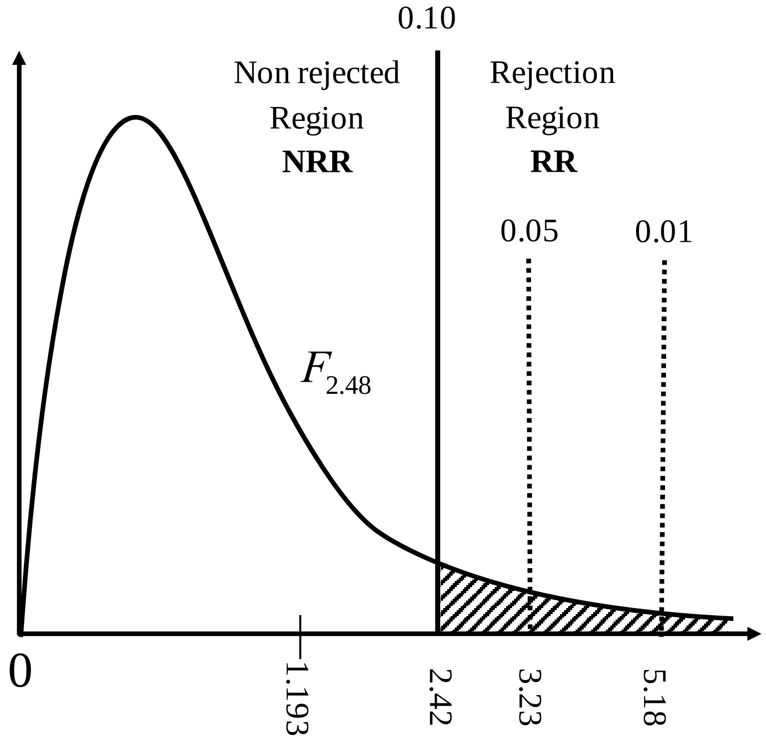

FIGURE 4.17. Example 4.10: Rejection region using *F* distribution (α values are from a $F_{2.40}$).

4.3 Testing multiple linear restrictions using the *F* test

EXAMPLE 4.11 Salaries of CEOs

$$\ln(salary) = \beta_1 + \beta_2 \ln(sales) + \beta_3 roe + \beta_4 ros + u$$

$$\widehat{\ln(salary_i)} = 4.3117 + 0.2803 \ln(sales_i) + 0.0174 roe_i + 0.00024 ros_i$$

$$R^2 = 0.283 \quad n = 209$$

$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$

$H_1 : H_0$ is not true

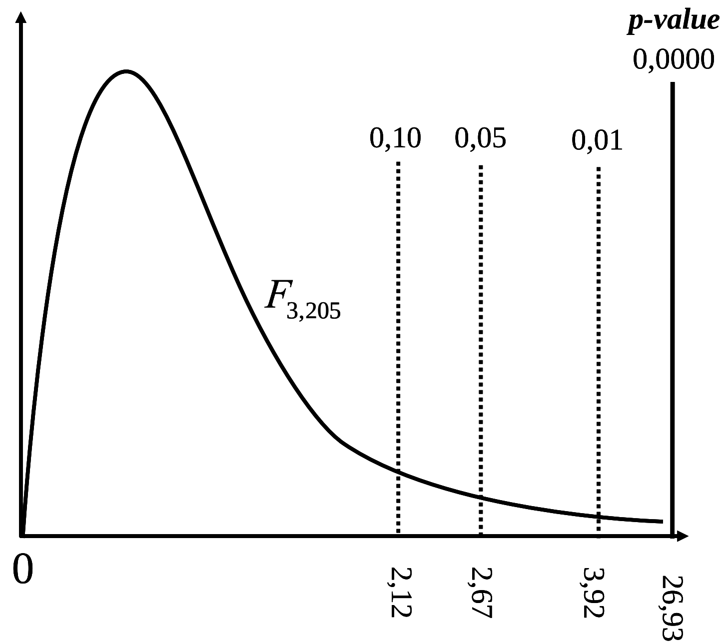

FIGURE 4.18. Example 4.11: ***p-value*** **using *F* distribution (*α* values are for a $F_{3,140}$)**

4.3 Testing multiple linear restrictions using the *F* test

EXAMPLE 4.11 Salaries of CEOs. (Continuation)

TABLE 4.9. Complete output from E-views in the example 4.11.

Dependent Variable: LOG(SALARY)
Method: Least Squares
Date: 04/12/12 Time: 19:39
Sample: 1 209
Included observations: 209

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.311712	0.315433	13.66919	0.0000
LOG(SALES)	0.280315	0.03532	7.936426	0.0000
ROE	0.017417	0.004092	4.255977	0.0000
ROS	0.000242	0.000542	0.446022	0.6561

R-squared	0.282685	Mean dependent var	6.950386
Adjusted R-squared	0.272188	S.D. dependent var	0.566374
S.E. of regression	0.483185	Akaike info criterion	1.402118
Sum squared resid	47.86082	Schwarz criterion	1.466086
Log likelihood	-142.5213	**F-statistic**	**26.9293**
Durbin-Watson stat	2.033496	**Prob(F-statistic)**	**0.0000**

4.3 Testing multiple linear restrictions using the *F* test

EXAMPLE 4.12 An additional restriction in the production function. (Continuation of example 4.7)

$$\ln(output) = \beta_1 + \beta_2 \ln(labor) + \beta_3 \ln(capital) + u \qquad RSS_{UR} = 0.8516$$

$$H_0 : \begin{cases} \beta_2 + \beta_3 = 1 \\ \beta_1 = 0 \end{cases}$$

$$H_1 : H_0 \text{ is not true}$$

$$\ln(output) = (1 - \beta_3)\ln(labor) + \beta_3 \ln(capital) + u$$

$$\ln(output / labor) = \beta_3 \ln(capital / labor) + u \qquad RSS_R = 3.1101$$

$$F = \frac{\left(RSS_R - RSS_{UR}\right)/q}{RSS_{UR}/(n-k)} = \frac{(3.1101 - 0.8516)/2}{0.8516/(27-3)} = 13.551$$

4.3 Testing multiple linear restrictions using the *F* test

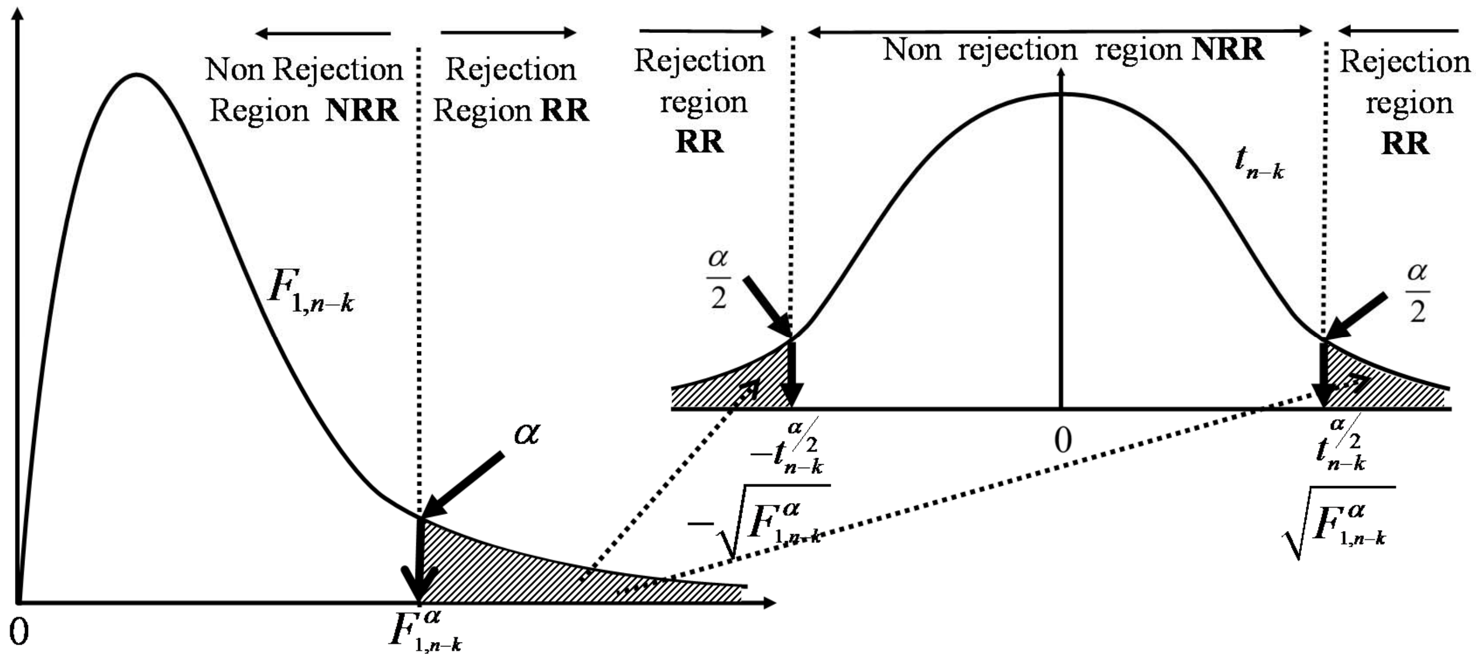

FIGURE 4.19. Relationship between a $F_{1,n-k}$ and t_{n-k}.

4.5 Prediction

EXAMPLE 4. 13 What is the expected score in the final exam with 7 marks in the first short exam?

Fitted model: $\widehat{finalmrk}_i = \underset{(0.715)}{4.155} + \underset{(0.123)}{0.491} shortex1_i \qquad \hat{\sigma} = 1.649 \quad R^2 = 0.533 \quad n = 16$

Fitted model with the regressor *shortex1*º=7:

$$\widehat{finalmrk}_i = \underset{(0.497)}{7.593} + \underset{(0.123)}{0.491} \left[shortex1_i - 7 \right] \qquad \hat{\sigma} = 1.649 \quad R^2 = 0.533 \quad n = 16$$

The point prediction for *shortex*1º=7: $\hat{\theta}_0$ =7.593

The lower and upper bounds of a 95% *CI*:

$$\underline{\theta}^0 = \hat{\theta}^0 - se(\hat{\theta}^0) \times t_{14}^{0.05/2} = 7.593 - 0.497 \times 2.14 = 6.5$$

$$\overline{\theta}^0 = \hat{\theta}^0 + se(\hat{\theta}^0) \times t_{14}^{0.05/2} = 7.593 + 0.497 \times 2.14 = 8.7$$

The point prediction by an alternative way: $\widehat{finalmrk} = 4.155 + 0.491 \times 7 = 7.593$

Estimation of the *se* of $\hat{e}_2^0$ $\qquad se(\hat{e}_2^0) = \left\{ \left[se(\hat{y}^0) \right]^2 + \hat{\sigma}^2 \right\}^{\frac{1}{2}} = \sqrt{0.497^2 + 1.649^2} = 1.722$

where 1.649 is the "S. E. of regression" obtained from the E-views output directly.
The lower and upper bounds of a 95% *probability interval*:

$$\underline{y}^0 = \hat{y}^0 - se(\hat{e}_2^0) \times t_{14}^{0.025} = 7.593 - 1.722 \times 2.14 = 3.7$$

$$\overline{y}^0 = \hat{y}^0 + se(\hat{e}_2^0) \times t_{14}^{0.025} = 7.593 + 1.722 \times 2.14 = 11.3$$

4.5 Prediction

EXAMPLE 4.14 Predicting the salary of CEOs

$$\widehat{salary}_i = \underset{(104)}{1381} + \underset{(0.0013)}{0.008377}\, assets_i + \underset{(8.671)}{32.508}\, tenure_i + \underset{(0.0538)}{0.2352}\, profits_i$$

$$\hat{\sigma} = 1506 \quad R^2 = 0.2404 \quad n = 447$$

TABLE 4.10. Descriptive measures of variables of the model on CEOs salary.

	assets	*tenure*	*profits*
Mean	27054	7.8	700
Median	7811	5.0	333
Maximum	668641	60.0	22071
Minimum	718	0.0	-2669
Observations	447	447	447

TABLE 4.11. Predictions for selected values.

	Prediction $\hat{\theta}_0$	*Std. Error* $se(\hat{\theta}_0)$
Mean values	2026	71
Median value	1688	78
Maximum values	14124	1110
Minimum values	760	195

4.5 Prediction

EXAMPLE 4.15 Predicting the salary of CEOs with a log model (continuation 4.14)

$$\widehat{\ln(salary_i)} = \underset{(0.210)}{5.5168} + \underset{(0.0232)}{0.1885}\ln(assets_i) + \underset{(0.0032)}{0.0125}\, tenure_i + \underset{(0.0000195)}{0.00007}\, profits_i$$

$$\hat{\sigma} = 0.5499 \quad R^2 = 0.2608 \quad n = 447$$

Inconsistent prediction

$$\widetilde{salary}_i = \exp(\widehat{\ln(salary_i)})$$

$$= \exp(5.5168 + 0.1885\ln(10000) + 0.0125\times 10 + 0.00007\times 1000) = 1207$$

Consistent prediction

$$\widehat{salary} = \exp(0.5499^2 / 2)\times 1207 = 1404$$