L11MEE TUTORIALS - SEMESTER 2 – TUTORIAL 1

By ERNESTO M. GAVASSA PEREZ*

---

*This document was compiled based on the answers provided by Professor Patrick Marsh. Hence, all the credit should be given to him

1. Let the continuous random variable (RV) X have the probability density function (PDF),

$$\boldsymbol{f(x) = cx^2, \qquad 0 \le x \le 3}$$

a) **Determine the value of c such that $f(x)$ represents a valid PDF.**

As discussed in the lectures, $f(x)$ represents a valid PDF if two conditions hold: **i)** $\boldsymbol{f(x) \ge 0}$ and **ii)** $\boldsymbol{\int_{-\infty}^{\infty} f(x)\, dx = 1}$**.** As $x^2 \ge 0$ for all values of x, then $c \ge 0$ for the first condition to hold. We can look more closely at the second condition by substituting $f(x)$ by cx^2 in the second condition and solving the integral. From the rules of calculus, we know that:

$$\boldsymbol{\int c * g(x) dx = c * \int g(x) dx}$$

Hence, we can rewrite the second condition as:

$$\int_0^3 c * x^2\, dx = c * \int_0^3 x^2\, dx = 1$$

Noticing that the range of x defines the bounds of the integral. In addition, from the rules of calculus we know that:

$$\boldsymbol{\int x^a dx = \frac{x^{a+1}}{a+1} + C}$$

Hence, we can rewrite the previous equation as:

$$c * \left[\frac{x^3}{3}\right]_0^3 = 1$$

The First Fundamental Theorem of Calculus states the following:

$$\boldsymbol{if\ f\ is\ continuous\ and\ F(x) = \int f(x) dx, \quad then \qquad \int_b^a f(x)\, dx = [F(x)]_b^a}$$
$$\boldsymbol{= F(x = a) - F(x = b)}$$

Applying it to our definite integral, we get:

$$c * \left(\frac{3^3}{3} - \frac{0^3}{3}\right) = c * (3^2 - 0) = 9 * c = 1$$

Isolating c, we get:

$$\boldsymbol{c = \frac{1}{9}}$$

b) Draw the PDF of X.

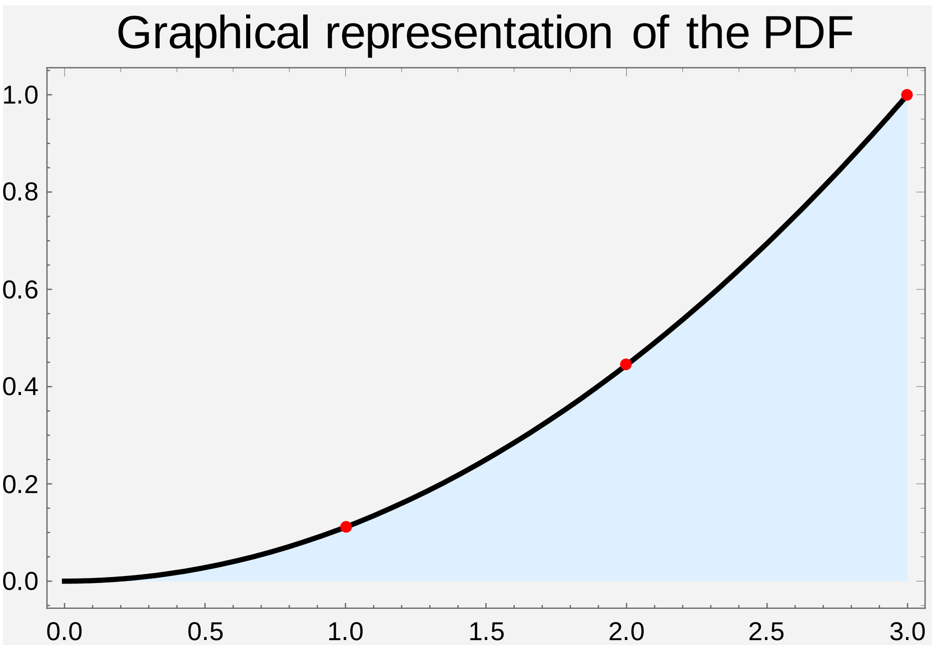
In order to draw the PDF, we first write it down:

$$y = \frac{x^2}{9}, for\ 0 \le x \le 3\ and\ 0 \le y \le 1$$

We, then, find some values of y for given values of x:

y	**x**
0	0
$\frac{1}{9}$	1
$\frac{4}{9}$	2
1	3

Finally, we plot those values. I have used Mathematica, a software widely used for mathematic applications. You can access this software from university computers.



c) Find the following probabilities and illustrate them graphically;

We have previously computed that $c * \int_b^a x^2\,dx = c * \left[\frac{x^3}{3}\right]_b^a$. As we now know that $c = \frac{1}{9}$, in the following exercises we directly use the formula $\frac{1}{9} * \left[\frac{x^3}{3}\right]_b^a$. We use the First Fundamental Theorem of Calculus, as stated above, to find the solutions.

a. $\mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{1}]$

$$\mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{1}] = \frac{1}{9} * \left[\frac{x^3}{3}\right]_{\mathbf{0}}^{\mathbf{1}} = \frac{1}{9} * \left[\frac{\mathbf{1}^3}{3} - \frac{\mathbf{0}^3}{3}\right] = \frac{1}{9} * \left[\frac{1}{3} - 0\right] = \frac{1}{9} * \left[\frac{1}{3}\right] = \frac{\mathbf{1}}{\mathbf{27}}$$

b. $\mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{2}]$

$$\mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{2}] = \frac{1}{9} * \left[\frac{x^3}{3}\right]_{\mathbf{0}}^{\mathbf{2}} = \frac{1}{9} * \left[\frac{\mathbf{2}^3}{3} - \frac{\mathbf{0}^3}{3}\right] = \frac{1}{9} * \left[\frac{8}{3} - 0\right] = \frac{1}{9} * \left[\frac{8}{3}\right] = \frac{\mathbf{8}}{\mathbf{27}}$$

c. $\mathbf{Pr}[\mathbf{1} \leq \boldsymbol{X} \leq \mathbf{2}]$

$$\mathbf{Pr}[\mathbf{1} \leq \boldsymbol{X} \leq \mathbf{2}] = \frac{1}{9} * \left[\frac{x^3}{3}\right]_{1}^{\mathbf{2}} = \frac{1}{9} * \left[\frac{\mathbf{2}^3}{3} - \frac{\mathbf{1}^3}{3}\right] = \frac{1}{9} * \left[\frac{8}{3} - \frac{1}{3}\right] = \frac{1}{9} * \left[\frac{7}{3}\right] = \frac{\mathbf{7}}{\mathbf{27}}$$

<u>**Extra comment**</u>: It is interesting to see that

$$\mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{2}] = \mathbf{Pr}[\mathbf{0} \leq \boldsymbol{X} \leq \mathbf{1}] + \mathbf{Pr}[\mathbf{1} \leq \boldsymbol{X} \leq \mathbf{2}]$$

But, why is this the case? If we abstract from the concepts we are dealing with (PDF's), we are just computing areas. In this case, the areas between two numbers just represent the probability that the value of a variable lies between two given numbers. Hence, it makes sense that the probability that a variable lies between 0 and 2 is just the sum of the probabilities that such a variable lies between 0 and 1 and the probability that our variable lies between 1 and 2.

2. Let the RVs X and Y have joint PDF

$$\boldsymbol{f(x,y) = 2 - x - y, \quad 0 \le x \le 1, \quad 0 \le y \le 1.}$$

a) Verify that $f(x,y)$ represents a valid PDF.

For a joint PDF to be valid, we need conditions analogous to the ones outlined in the first answer. Firstly, we want $f(x,y) \ge 0 \; for \; 0 \le x \le 1, \;\; 0 \le y \le 1$ to hold for every possible value in the range of x and y. In order to verify that this holds true, we compute the value of $f(x,y)$ at the least upper bound of x and y. We define a least upper bound as the highest value that x and y can take that lies within their corresponding ranges. In our case, the least upper bounds of x and y, labelled as $\sup(X)$ and $\sup(Y)$ respectively, are $\sup(X) = 1$ and $\sup(Y) = 1$. Hence,

$$f(x = \sup(X), y = \sup(Y)) = 2 - \sup(X) - \sup(Y) = 2 - 1 - 1 = 0 \ge 0$$

As we can see, the first condition for the existence of a PDF is satisfied for the proposed range of values of x and y. Notice that this wouldn't hold true, for instance, for $0 \le x \le 2$; since $f(x = 2, y = 1) = 2 - 2 - 1 = -1 \le 0$.

Secondly, we want that the area of the joint PDF between the ranges of x and y is equal to 1, meaning that the probability that the random variables take a number between all the proposed values can never be greater than certainty. More formally,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \boldsymbol{f(x,y)\, dx\, dy = 1}$$

Were $-\infty$ and ∞ refer to generic upper and lower bounds of x and y. As, in our case, $0 \le x \le 1$, $0 \le y \le 1$ and $f(x,y) = 2 - x - y$, then the condition becomes:

$$\int_0^1 \int_0^1 (2 - x - y)\, dx\, dy = 1$$

Whenever we have two integrals with respect to two different variables, we proceed in the following way. First, we integrate with respect to one variable, leaving the other one as a constant. Secondly, we integrate with respect to the remaining variable. As we are dealing with definite integrals, after the first and second integration we need to use the First Fundamental Theorem of Calculus. From the rules of calculus, we know that:

$$\int \boldsymbol{b} + \boldsymbol{x}^{\boldsymbol{a}}\, \boldsymbol{dx} = \boldsymbol{b} * \boldsymbol{x} + \frac{\boldsymbol{x}^{\boldsymbol{a}+\boldsymbol{1}}}{\boldsymbol{a}+\boldsymbol{1}} + \boldsymbol{C}$$

We start by integrating with respect to x. Hence, y is just a constant and should be treated as the "b" in the rule above:

$$\int_0^1 \left[2 * x - y * x - \frac{x^2}{2}\right]_{\mathbf{0}}^{\mathbf{1}} dy = 1$$

Using the First Fundamental Theorem of Calculus, we get:

$$\int_0^1 \left\{\left(2 * (\mathbf{1}) - y * (\mathbf{1}) - \frac{(\mathbf{1})^2}{2}\right) - \left(2 * (\mathbf{0}) - y * (\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right)\right\} dy = 1$$

Which can be simplified to get the following equation:

$$\int_0^1 \left(2 - y - \frac{1}{2}\right) dy = \int_0^1 \left(\frac{3}{2} - y\right) dy = 1$$

We know integrate with respect to y to get:

$$\left[\frac{3}{2} * y - \frac{y^2}{2}\right]_{\mathbf{0}}^{\mathbf{1}} = 1$$

Now, using the First Fundamental Theorem of Calculus, we get:

$$\left[\frac{3}{2}(\mathbf{1}) - \frac{(\mathbf{1})^2}{2}\right] - \left[\frac{3}{2}(\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right] = \frac{3}{2} - \frac{1}{2} = \frac{2}{2} = \mathbf{1} = \mathbf{1}$$

Hence, we can see that the second condition holds, as well, true for the proposed ranges of x and y.

b) Determine the marginal density functions of X and Y.

In order to find the marginal density functions of X and Y, we need to integrate $f(x, y)$ with respect to the opposite variable for which we want to find the marginal density. For x, this implies:

$$f_X(x) = \int_0^1 (2 - x - y)\, dy$$

Using the rules described above, we get:

$$f_X(x) = \left[2 * y - x * y - \frac{y^2}{2}\right]_{\mathbf{0}}^{\mathbf{1}}$$

$$= \left[2 * (\mathbf{1}) - x * (\mathbf{1}) - \frac{(1)^2}{2}\right] - \left[2 * (\mathbf{0}) - x * (\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right]$$

Which simplifies to:

$$\boldsymbol{f_X(x) = \left[2 - x - \frac{1}{2}\right] = \frac{3}{2} - x}$$

Proceeding analogously for y, we have the following condition:

$$f_Y(y) = \int_0^1 (2 - x - y)\, dx$$

Using the rules described above, we get:

$$f_Y(y) = \left[2 * x - y * x - \frac{x^2}{2}\right]_{\mathbf{0}}^{\mathbf{1}}$$

$$= \left[2 * (\mathbf{1}) - y * (\mathbf{1}) - \frac{(1)^2}{2}\right] - \left[2 * (\mathbf{0}) - y * (\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right]$$

Which simplifies to:

$$\boldsymbol{f_Y(y) = \left[2 - y - \frac{1}{2}\right] = \frac{3}{2} - y}$$

c) **Evaluate the following probabilities,**

This section is a little bit different from the analogous one in the first exercise, as it implies dealing with joint density functions. Here, when we get asked $\Pr[0 \leq X \leq a]$ or $\Pr[0 \leq Y \leq a]$, we make reference to the marginal cumulative probabilities. This is because we want to isolate the influence of the variable we are interested in, and we do so by analysing the probabilities in the full range of the other variables. This implies that one needs to use the marginal density functions found in the exercise above. However, when we are asked to find $\Pr[0 \leq X \leq a,\ \ 0 \leq Y \leq b\ ]$, we need to compute the joint cumulative probability given the ranges of x and y proposed.

a. $\mathbf{Pr}\left[\mathbf{0} \leq \boldsymbol{X} \leq \frac{\mathbf{1}}{\mathbf{2}}\right]$

Given the proposed range of x, we need to compute the definite integral of the marginal density of x described below:

$$\Pr\left[0 \leq X \leq \frac{1}{2}\right] = \int_{0}^{\frac{1}{2}} \left(\frac{3}{2} - x\right) dx$$

Applying the rules mentioned above, we get:

$$\Pr\left[0 \leq X \leq \frac{1}{2}\right] = \left[\frac{3}{2} * x - \frac{x^2}{2}\right]_{\mathbf{0}}^{\frac{\mathbf{1}}{\mathbf{2}}}$$

Which, using the First Fundamental Theorem of Calculus, becomes:

$$\mathbf{Pr}\left[\mathbf{0} \leq \boldsymbol{X} \leq \frac{\mathbf{1}}{\mathbf{2}}\right] = \left[\frac{3}{2} * \left(\frac{\mathbf{1}}{\mathbf{2}}\right) - \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2}\right] - \left[\frac{3}{2} * (\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right] = \left[\frac{3}{4} - \frac{\frac{1}{4}}{2}\right] = \frac{3}{4} - \frac{1}{8} = \frac{6-1}{8}$$
$$= \frac{\mathbf{5}}{\mathbf{8}}$$

b. $\mathbf{Pr}\left[\mathbf{0} \leq \boldsymbol{X} \leq \frac{\mathbf{1}}{\mathbf{2}},\ \ \mathbf{0} \leq \boldsymbol{Y} \leq \frac{\mathbf{1}}{\mathbf{2}}\right]$

As explained before, a range for each variable implies a joint cumulative probability. This can be computed, given the proposed ranges, with the following integral:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right] = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} (2 - x - y)\, dx\, dy$$

As explained in part a), we first integrate with one variable, treating the other one as constant. Afterwards, we apply the first Fundamental Theorem of Calculus and repeat those two steps. We start by integrating with respect to x. Following the aforementioned rules, we get:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right] = \int_0^{\frac{1}{2}} \left[2 * x - y * x - \frac{x^2}{2}\right]_{\mathbf{0}}^{\mathbf{\frac{1}{2}}} dy$$

Applying the First fundamental Theorem of Calculus, we get:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right]$$
$$= \int_0^{\frac{1}{2}} \left\{\left[2 * \left(\frac{\mathbf{1}}{\mathbf{2}}\right) - y * \left(\frac{\mathbf{1}}{\mathbf{2}}\right) - \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2}\right] - \left[2 * (\mathbf{0}) - y * (\mathbf{0}) - \frac{(\mathbf{0})^2}{2}\right]\right\} dy$$

Which, after simplifying, becomes:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right] = \int_0^{\frac{1}{2}} \left\{\left[1 - \frac{y}{2} - \frac{\frac{1}{4}}{2}\right]\right\} dy = \int_0^{\frac{1}{2}} \left\{\left[1 - \frac{y}{2} - \frac{1}{8}\right]\right\} dy$$
$$= \int_0^{\frac{1}{2}} \left(\frac{7}{8} - \frac{y}{2}\right) dy$$

If we now integrate with respect to y, we get:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right] = \left[\frac{7}{8} * y - \frac{1}{2} * \frac{y^2}{2}\right]_{\mathbf{0}}^{\mathbf{\frac{1}{2}}}$$

Again, applying the First Fundamental Theorem of Calculus, we get:

$$\Pr\left[0 \le X \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\right] = \left[\frac{7}{8} * \left(\frac{\mathbf{1}}{\mathbf{2}}\right) - \frac{1}{2} * \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2}\right] - \left[\frac{7}{8} * (\mathbf{0}) - \frac{1}{2} * \frac{(\mathbf{0})^2}{2}\right]$$

Which simplifies to:

$$\mathbf{Pr}\left[\mathbf{0} \leq \boldsymbol{X} \leq \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0} \leq \boldsymbol{y} \leq \frac{\mathbf{1}}{\mathbf{2}}\right] = \left[\frac{7}{16} - \frac{1}{2} * \frac{\frac{1}{4}}{2}\right] = \frac{7}{16} - \frac{1}{2} * \frac{1}{8} = \frac{7}{16} - \frac{1}{16} = \frac{6}{16} = \frac{\mathbf{3}}{\mathbf{8}}$$

3. Let the RVs X and Y have joint PDF

$$\boldsymbol{f(x,y) = x^2 + \frac{x*y}{3}, \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 2.}$$

a) Are X and Y statistically independent?

In order to check whether X and Y are statistically independent, we need to compare the joint PDF with the product of the marginal densities of X and Y. They will be statistically independent if the following relation holds:

$$f_X(x) * f_Y(y) = f(x,y)$$

As the joint PDF is already given in the question, we next step is to compute the marginal densities and, afterwards, compute their product. Following the logic of the previous question, the marginal density of X is given by:

$$f_X(x) = \int_0^2 \left(x^2 + \frac{x*y}{3}\right) dy$$

Using the rules outlined above, we get:

$$f_X(x) = \left[x^2 * y + \frac{x}{3} * \frac{y^2}{2}\right]_{\mathbf{0}}^{\mathbf{2}} = \left[x^2 * (\mathbf{2}) + \frac{x}{3} * \frac{(\mathbf{2})^2}{2}\right] - \left[x^2 * (\mathbf{0}) + \frac{x}{3} * \frac{(\mathbf{0})^2}{2}\right]$$

Which simplifies to:

$$f_X(x) = 2 * x^2 + \frac{4 * x}{6} = 2 * x^2 + \frac{2 * x}{3}$$

The marginal density of Y is given by:

$$f_Y(y) = \int_0^1 \left(x^2 + \frac{x*y}{3}\right) dx$$

Using the rules outlined above, we get:

$$f_Y(y) = \left[\frac{x^3}{3} + \frac{x^2}{2} * \frac{y}{3}\right]_{\mathbf{0}}^{\mathbf{1}} = \left[\frac{(\mathbf{1})^3}{3} + \frac{(\mathbf{1})^2}{2} * \frac{y}{3}\right] - \left[\frac{(\mathbf{0})^3}{3} + \frac{(\mathbf{0})^2}{2} * \frac{y}{3}\right]$$

Which simplifies to:

$$f_Y(y) = \frac{1}{3} + \frac{y}{6}$$

Multiplying both marginal density function yields:

$$f_X(x) * f_Y(y) = \left(2 * x^2 + \frac{2 * x}{3}\right) * \left(\frac{1}{3} + \frac{y}{6}\right)$$

Expanding the Right Hand Side of the previous equation, we get:

$$f_X(x) * f_Y(y) = \frac{2 * x^2}{3} + \frac{2 * x^2 * y}{6} + \frac{2 * x}{9} + \frac{2 * x * y}{18}$$

Which simplifies to:

$$f_X(x) * f_Y(y) = \frac{x^2 * (2 + y)}{3} + \frac{x * (2 + y)}{9}$$

Comparing the previous equation with the joint PDF, it is clear to see that the condition for statistical independence does not hold for all the possible values of x and y*:

$$\boldsymbol{\frac{x^2 * (2 + y)}{3} + \frac{x * (2 + y)}{9} \neq x^2 + \frac{x * y}{3}}$$

* It does hold only if $x = \frac{2}{3}$. However, note that this is only one point in the whole range of values that x can take. As far as the equality does not hold for all the possible values of x and y, then the two variables are said to be dependent.

b) Evaluate the following probabilities,

The first to subsections are straightforward and follow the same logic as the ones in the previous exercise. The last two subsections follow 1.5 in the lecture notes closely. In order to find $Pr[(0 \leq Y \leq a|X = b)]$ one needs to divide the joint PDF by the marginal pdf of X, both of them evaluated at the relevant ranges.

a. $\boldsymbol{Pr\left[X \geq \frac{1}{2}\right]}$

We need to use the marginal PDF of X and find the definite integral between $\frac{1}{2}$ and 1:

$$\Pr\left[X \geq \frac{1}{2}\right] = \int_{\frac{1}{2}}^{1} f_X(x)\,dx = \int_{\frac{1}{2}}^{1} 2 * x^2 + \frac{2 * x}{3} dx$$

Using the rules outlined above, we get:

$$\Pr\left[X \geq \frac{1}{2}\right] = \int_{\frac{1}{2}}^{1} f_X(x)\,dx = \left[\frac{2 * x^3}{3} + \frac{2}{3} * \frac{x^2}{2}\right]_{\frac{1}{2}}^{1}$$

Using the First Fundamental Theorem of Calculus, we get:

$$\Pr\left[X \geq \frac{1}{2}\right] = \left[\frac{2 * (1)^3}{3} + \frac{2}{3} * \frac{(1)^2}{2}\right] - \left[\frac{2 * \left(\frac{1}{2}\right)^3}{3} + \frac{2}{3} * \frac{\left(\frac{1}{2}\right)^2}{2}\right]$$

Which, simplifying, is equal to:

$$\mathbf{Pr}\left[\boldsymbol{X} \geq \frac{\mathbf{1}}{\mathbf{2}}\right] = \left[\frac{2}{3} + \frac{2}{6}\right] - \left[\frac{2 * \frac{1}{8}}{3} + \frac{2}{3} * \frac{\frac{1}{4}}{2}\right] = \left[\frac{2}{3} + \frac{1}{3}\right] - \left[\frac{1}{12} + \frac{1}{12}\right] = 1 - \frac{2}{12} = \frac{10}{12} = \frac{\mathbf{5}}{\mathbf{6}}$$

b. $\boldsymbol{Pr\left[0 \leq Y \leq \frac{1}{2},\quad 0 \leq X \leq \frac{1}{2}\right]}$

We need to find the definite integral of the joint PDF at the relevant ranges for X and Y:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} \left(x^2 + \frac{x * y}{3}\right) dxdy$$

As we did in the previous exercise, we need to integrate with respect to one variable while treating the other one as a constant, use the First Fundamental Theorem of Calculus and repeat with respect to the remaining variable. We start by integrating with respect to x:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \int_{0}^{\frac{1}{2}} \left[\frac{x^3}{3} + \frac{x^2}{2} * \frac{y}{3}\right]_{\mathbf{0}}^{\frac{\mathbf{1}}{\mathbf{2}}} dy$$

Applying the First fundamental Theorem of Calculus, we get:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \int_{0}^{\frac{1}{2}} \left\{\left[\frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^3}{3} + \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2} * \frac{y}{3}\right] - \left[\frac{(\mathbf{0})^3}{3} + \frac{(\mathbf{0})^2}{2} * \frac{y}{3}\right]\right\} dy$$

Which, by simplifying, is equal to:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \int_{0}^{\frac{1}{2}} \left\{\left[\frac{\frac{1}{8}}{3} + \frac{\frac{1}{4}}{2} * \frac{y}{3}\right]\right\} dy = \int_{0}^{\frac{1}{2}} \left\{\left[\frac{1}{24} + \frac{y}{24}\right]\right\} dy$$

Now, integrating with respect to y yields:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \left[\frac{y}{24} + \frac{1}{24} * \frac{y^2}{2}\right]_{\mathbf{0}}^{\frac{\mathbf{1}}{\mathbf{2}}}$$

Again, by using the First Fundamental Theorem of Calculus, we get:

$$\Pr\left[0 \leq Y \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right] = \left[\frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)}{24} + \frac{1}{24} * \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2}\right] - \left[\frac{(\mathbf{0})}{24} + \frac{1}{24} * \frac{(\mathbf{0})^2}{2}\right]$$

Which, when simplifying, becomes:

$$\mathbf{Pr}\left[\mathbf{0} \leq \boldsymbol{Y} \leq \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0} \leq \boldsymbol{X} \leq \frac{\mathbf{1}}{\mathbf{2}}\right] = \frac{1}{48} + \frac{1}{24} * \frac{\frac{1}{4}}{2} = \frac{1}{48} + \frac{1}{24} * \frac{1}{8} = \frac{4}{192} + \frac{1}{192} = \frac{\mathbf{5}}{\mathbf{192}}$$

c. $\boldsymbol{Pr}\left[\left(\mathbf{0} \leq \boldsymbol{Y} \leq \frac{\mathbf{1}}{\mathbf{2}} \middle| \boldsymbol{X} = \frac{\mathbf{1}}{\mathbf{2}}\right)\right]$

As stated before, we need to divide the joint PDF by the marginal PDF X, both of them evaluated at the relevant ranges. As $X = \frac{1}{2}$, we need to find $f_X(x = \frac{1}{2})$:

$$f_X\left(x=\frac{1}{2}\right)=2*\left(\frac{1}{2}\right)^2+\frac{2*\left(\frac{1}{2}\right)}{3}$$

Which is equivalent to:

$$f_X\left(x=\frac{1}{2}\right)=2*\frac{1}{4}+\frac{1}{3}=\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$$

Hence,

$$Pr\left[\left(0\le Y\le\frac{1}{2}\middle|X=\frac{1}{2}\right)\right]=\frac{\int_0^{\frac{1}{2}}f\left(x=\frac{1}{2},y\right)dy}{f_X\left(x=\frac{1}{2}\right)}=\frac{\int_0^{\frac{1}{2}}\left(\left(\frac{1}{2}\right)^2+\frac{\left(\frac{1}{2}\right)*y}{3}\right)dy}{\frac{5}{6}}$$

Which is equivalent to:

$$Pr\left[\left(0\le Y\le\frac{1}{2}\middle|X=\frac{1}{2}\right)\right]=\frac{\int_0^{\frac{1}{2}}\left(\frac{1}{4}+\frac{y}{6}\right)dy}{\frac{5}{6}}$$

Integrating with respect to y, we get:

$$Pr\left[\left(0\le Y\le\frac{1}{2}\middle|X=\frac{1}{2}\right)\right]=\frac{\left[\frac{y}{4}+\frac{1}{6}*\frac{y^2}{2}\right]_{\mathbf{0}}^{\mathbf{\frac{1}{2}}}}{\frac{5}{6}}$$

Applying the First Fundamental Theorem of Calculus, we get:

$$Pr\left[\left(0\le Y\le\frac{1}{2}\middle|X=\frac{1}{2}\right)\right]=\frac{\left[\frac{\left(\mathbf{\frac{1}{2}}\right)}{4}+\frac{1}{6}*\frac{\left(\mathbf{\frac{1}{2}}\right)^2}{2}\right]-\left[\frac{\mathbf{0}}{4}+\frac{1}{6}*\frac{\mathbf{0}^2}{2}\right]}{\frac{5}{6}}$$

Simplifying, we get:

$$\boldsymbol{Pr\left[\left(0\le Y\le\frac{1}{2}\middle|X=\frac{1}{2}\right)\right]}=\frac{\left[\frac{1}{8}+\frac{1}{6}*\frac{\frac{1}{4}}{2}\right]}{\frac{5}{6}}=\frac{\left[\frac{1}{8}+\frac{1}{6}*\frac{1}{8}\right]}{\frac{5}{6}}=\frac{\left[\frac{6+1}{48}\right]}{\frac{5}{6}}=\frac{7}{48}*\frac{6}{5}=\frac{7*6}{6*8*5}$$

$$=\mathbf{\frac{7}{40}}$$

d. $\boldsymbol{Pr\left[\left(0 \le X \le \frac{1}{2} \middle| Y = \frac{1}{2}\right)\right]}$

As $Y = \frac{1}{2}$, we need to find $f_Y(y = \frac{1}{2})$:

$$f_Y\left(y = \frac{1}{2}\right) = \frac{1}{3} + \frac{\frac{1}{2}}{6} = \frac{1}{3} + \frac{1}{12} = \frac{5}{12}$$

Now, we apply the formula discussed before:

$$Pr\left[\left(0 \le X \le \frac{1}{2} \middle| Y = \frac{1}{2}\right)\right] = \frac{\int_0^{\frac{1}{2}} f\left(x, y = \frac{1}{2}\right) dx}{f_Y\left(y = \frac{1}{2}\right)} = \frac{\int_0^{\frac{1}{2}} \left(x^2 + \frac{\left(\frac{1}{2}\right) * x}{3}\right) dx}{\frac{5}{12}} = \frac{\int_0^{\frac{1}{2}} \left(x^2 + \frac{x}{6}\right) dx}{\frac{5}{12}}$$

Integrating with respect to x, we get:

$$Pr\left[\left(0 \le X \le \frac{1}{2} \middle| Y = \frac{1}{2}\right)\right] = \frac{\left[\frac{x^3}{3} + \frac{1}{6} * \frac{x^2}{2}\right]_{\mathbf{0}}^{\frac{\mathbf{1}}{\mathbf{2}}}}{\frac{5}{12}}$$

Applying the First Fundamental Theorem of Calculus, we get:

$$Pr\left[\left(0 \le X \le \frac{1}{2} \middle| Y = \frac{1}{2}\right)\right] = \frac{\left[\frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^3}{3} + \frac{1}{6} * \frac{\left(\frac{\mathbf{1}}{\mathbf{2}}\right)^2}{2}\right] - \left[\frac{(0)^3}{3} + \frac{1}{6} * \frac{(\mathbf{0})^2}{2}\right]}{\frac{5}{12}}$$

Simplifying, we get:

$$\boldsymbol{Pr\left[\left(0 \le X \le \frac{1}{2} \middle| Y = \frac{1}{2}\right)\right]} = \frac{\left[\frac{\frac{1}{8}}{3} + \frac{1}{6} * \frac{\frac{1}{4}}{2}\right]}{\frac{5}{12}} = \frac{\frac{1}{24} + \frac{1}{48}}{\frac{5}{12}} = \frac{\frac{2+1}{48}}{\frac{5}{12}} = \frac{3}{48} * \frac{12}{5} = \frac{3 * 2 * 6}{6 * 2 * 4 * 5}$$

$$= \boldsymbol{\frac{3}{20}}$$

L11MEE TUTORIALS – SEMESTER 2 – TUTORIAL 2

By ERNESTO M. GAVASSA PEREZ*

---

*This document was compiled based on the answers provided by Professor Patrick Marsh. Hence, all the credit should be given to him

Question 1. A firm's profit is made by combining sand and water. A unit of the product contains $\boldsymbol{X}$% of sand, where $\boldsymbol{X}$ is considered to be a RV. Suppose that $\boldsymbol{X}$ has the PDF

$$\boldsymbol{f(x) = \frac{3x(10-x)}{500}}, \qquad \boldsymbol{0 \le X \le 10}.$$

Let $\boldsymbol{p}$ denote the profit per unit, be the following function of $\boldsymbol{X}$,

$$\boldsymbol{p = a + bX}.$$

Compute the expected profit per unit.

First, note that the question already gives the profit: there is no need to further substract any cost! Secondly, to find the expected profit per unit we need to calculate the expected value of p, as defined below:

$$E[p] = E[a + bX]$$

Because of the rules of expectations outlined in the lecture notes, we know that $E[a] = a$ and $E[bX] = bE[X]$. Hence, we can rewrite the previous equation as:

$$E[p] = a + bE[X]$$

We know that the expected value of the RV X can be expressed as:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Substituting $E[X]$ in the equation provided for $E[p]$, we get:

$$E[p] = a + b \int_{-\infty}^{\infty} xf(x)dx$$

Substituting the generic bounds of the integral along with the specific functional form of $f(x)$ given our setup, we get:

$$E[p] = a + b\int_{0}^{10} x\frac{3x(10-x)}{500}dx$$

Expanding the term inside the integral, we get:

$$E[p] = a + b\int_{0}^{10} \frac{30x^2 - 3x^3}{500}dx$$

Using the rules of calculus (specifically, using the fact that $\int x^a dx = \frac{x^{a+1}}{a+1} + C$ and $\int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx$), we get:

$$E[p] = a + b\left[\frac{30}{500} * \frac{x^3}{3} - \frac{3}{500} * \frac{x^4}{4}\right]_{\mathbf{0}}^{\mathbf{10}}$$

Applying the First Fundamental Theorem of Calculus, we get:

$$E[p] = a + b\left\{\left[\frac{30}{500} * \frac{\mathbf{10}^3}{3} - \frac{3}{500} * \frac{\mathbf{10}^4}{4}\right] - \left[\frac{30}{500} * \frac{\mathbf{0}^3}{3} - \frac{3}{500} * \frac{\mathbf{0}^4}{4}\right]\right\}$$

Which, after simplifying, collapses to:

$$E[p] = a + b\left\{\frac{30}{500} * \frac{1{,}000}{3} - \frac{3}{500} * \frac{10{,}000}{4}\right\} = a + b[20 - 15] = a + 5b$$

Answer: Hence, the expected profit per unit of the firm is $\boldsymbol{a + 5b}$.

Question 2. An electronic device has a life length X (which is a RV) with the following PDF,

$$\boldsymbol{f(x) = e^{-x}, \qquad x > 0.}$$

Suppose that the cost of manufacturing such an item is £2. The manufacturer sells the item for £6, but guarantees a full refund if $X \leq 0.9$. What is the manufacturer's expected profit per item?

The first thing we need to do is to calculate the profit of the manufacturer in all the potential situations that may arise. If $X > 0.9$, the manufacturer earns £6 by selling the item and pays £2 for producing it, making a total profit of £4. If, in contrast, $X \leq 0.9$, then the manufacturer pays the production cost of £2 but refunds the full price of the electronic device to the client; making, thus, a loss of £2. We can represent these two situations with a step function, as shown below:

$$p = \begin{cases} £4 \quad if\ X > 0.9 \\ -£2\ if\ X \leq 0.9 \end{cases}$$

The expected profit of the manufacturer in a given state of nature is calculated by multiplying the profit of the manufacturer in a given state by the probability of such state of nature happening. The expected profit, $E[p]$, is calculated by aggregating the expected profit at each of the states of nature. In our case there are only two states of nature: either $X > 0.9$ or $X \leq 0.9$. Hence, the expected profit of the manufacturer can be written as:

$$E[p] = £4 * \Pr[X > 0.9] - 2£ * \Pr[X \leq 0.9]$$

We know that, in order for $f(x)$ to represent a valid PDF, then the following property needs to hold true:

$$\Pr[X \leq 0.9] + \Pr[X > 0.9] = \int_0^\infty x f(x)\, dx = 1$$

This is so because $x > 0$ represents all the potential values that x can take. Hence, the probability that x can take any of its potential values has to be equal to 1. Isolating $\Pr[X > 0.9]$ in the left hand side, we get:

$$\Pr[X > 0.9] = 1 - \Pr[X \leq 0.9]$$

We can use this equivalence to rewrite the expected profit as:

$$E[p] = £4 * (1 - \Pr[X \leq 0.9]) - 2£ * \Pr[X \leq 0.9]$$

Noting that $\Pr[X \leq x]$ defines the probability that X takes all values no bigger than x, we can use the PDF given in the question to get:

$$E[p] = £4 * \left(1 - \int_0^{0.9} e^{-x}\, dx\right) - 2£ * \left(\int_0^{0.9} e^{-x}\, dx\right)$$

By acknowledging $\frac{de^{f(x)}}{dx} = e^{f(x)} * f'(x)$ and $-1 = \frac{d(-x)}{dx}$, we can see that the following rule holds true: $-e^{-x} = \frac{de^{-x}}{dx}$. Hence,

$$E[p] = £4 * \left(1 - [-e^{-x}]_{\mathbf{0}}^{\mathbf{0.9}}\right) - 2£ * [-e^{-x}]_{\mathbf{0}}^{\mathbf{0.9}}$$

Applying the first Fundamental Theorem of Calculus, we get:

$$E[p] = £4 * (1 - \{[-e^{-\mathbf{0.9}}] - [-e^{-\mathbf{0}}]\}) - 2£ * \{[-e^{-\mathbf{0.9}}] - [-e^{-\mathbf{0}}]\}$$

As $e^{-0} = \frac{1}{e^0} = \frac{1}{1} = 1$, and $e^{-0.9} = \frac{1}{e^{0.9}} = \frac{1}{2.4596} = 0.4066$, the previous equation collapses to:

$$E[p] = £4 * (1 - \{-0.4066 + 1\}) - 2£ * \{-0.4066 + 1\}$$

As $-0.4066 + 1 = 0.5934$, we get:

$$E[p] = £4 * \left(1 - (0.5934)\right) - 2£ * (0.5934) = -£4 * (0.4066) - £2 * (0.5934) = £0.4394$$

Answer: Hence, the manufacturer's expected profit per item is $£0.4394$

Question 3. Let the RV X have PDF

$$\boldsymbol{f(x) = 1 - |x|, \quad -1 \leq X \leq 1,}$$

Where $|x|$ denotes the absolute value of x. Find both$E[X]$ and $V[X]$.

The absolute value function always reports a positive value, regardless of the sign of the original input. It is more intuitive to see it as a way to measure distance from the origin. Whenever the original input is positive the absolute value functions keeps the output the same as the input. However, if the original input is negative the absolute value function transforms the negative input into a positive input of the same magnitude. In order to transform a negative number into a positive one, we can use the rule $-(-x) = x$. Thus, we can rewrite the absolute value function in the following way:

$$|x| = \begin{cases} x \; if \; x \geq 0 \\ -x \; if \; x \leq 0 \end{cases}$$

As X can take both positive and negative values, the PDF will vary according to the sign of X. More specifically, $f(x) = 1 - x \; if \; x \geq 0$ and $f(x) = 1 - (-x) = 1 + x \; if \; x \leq 0$. We will use this later on. At the moment, let's just write the expected value of X:

$$E[X] = \int_{-1}^{1} x * (1 - |x|) \, dx$$

From the rules of calculus, we know that:

$$\int_{c}^{a} f(x) \, dx = \int_{b}^{a} f(x) \, dx + \int_{c}^{b} f(x) \, dx$$

Hence, it follows that:

$$E[X] = \int_{-1}^{1} x * (1 - |x|) \, dx = \int_{-1}^{0} x * (1 - |x|) \, dx + \int_{0}^{1} x * (1 - |x|) \, dx$$

Noticing that X takes negative values in $\int_{-1}^{0} x * (1 - |x|)\, dx$ and positive values in $\int_{0}^{1} x * (1 - |x|)\, dx$, we can substitute $(1 - |x|)$ by the two rules commented earlier to get:

$$E[X] = \int_{-1}^{0} x * (1 + x)\, dx + \int_{0}^{1} x * (1 - x)\, dx$$

Once we have gotten rid of the absolute value function, the calculations are straightforward and resemble the ones in previous exercises. Expanding the terms inside the integrals, we get:

$$E[X] = \int_{-1}^{0} x + x^2\, dx + \int_{0}^{1} x - x^2\, dx$$

Solving the integral, we get:

$$E[X] = \left[\frac{x^2}{2} + \frac{x^3}{3}\right]_{-1}^{0} + \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{0}^{1}$$

Applying the First fundamental Theorem of Calculus, we get:

$$E[X] = \left[\frac{\mathbf{0}^2}{2} + \frac{\mathbf{0}^3}{3}\right] - \left[+\frac{(\mathbf{-1})^2}{2} + \frac{(\mathbf{-1})^3}{3}\right] + \left[\frac{\mathbf{1}^2}{2} - \frac{\mathbf{1}^3}{3}\right] - \left[\frac{\mathbf{0}^2}{2} - \frac{\mathbf{0}^3}{3}\right]$$

Which, when simplifying, is equivalent to:

$$E[X] = -\left[\frac{1}{2} - \frac{1}{3}\right] + \left[\frac{1}{2} - \frac{1}{3}\right] = -\left[\frac{3}{6} - \frac{2}{6}\right] + \left[\frac{3}{6} - \frac{2}{6}\right] = \frac{1}{6} - \frac{1}{6} = 0$$

We know that the variance is equal to the square of the expected value around its mean. More formally,

$$V[X] = E[(X - E[X])^2]$$

However, as we know that $E[X] = 0$, the former expression becomes:

$$V[X] = E[(X - 0)^2] = E[X^2]$$

From the lecture notes, we know that we can rewrite such an expression as:

$$V[X] = E[X^2] = \int_{-1}^{1} x^2 f(x)\,dx = \int_{-1}^{1} x^2(1-|x|)\,dx$$

We deal with the absolute value in the same manner as before. Hence, our expression for the variance becomes:

$$V[X] = \int_{-1}^{0} x^2(1+x)\,dx + \int_{0}^{1} x^2(1-x)\,dx$$

Expanding the inside of the integrals, we get:

$$V[X] = \int_{-1}^{0} \frac{x^2}{2} + \frac{x^3}{3}dx + \int_{0}^{1} x^2 - \frac{x^3}{3}dx$$

Using the rules of calculus, we can solve the two integrals to get the following expression:

$$V[X] = \left[\frac{x^3}{3} + \frac{x^4}{4}\right]_{\mathbf{-1}}^{\mathbf{0}} + \left[\frac{x^3}{3} - \frac{x^4}{4}\right]_{\mathbf{0}}^{\mathbf{1}}$$

Using the First Fundamental Theorem of Calculus, we get:

$$V[X] = \left[\frac{\mathbf{0}^3}{3} + \frac{\mathbf{0}^4}{4}\right] - \left[\frac{(\mathbf{-1})^3}{3} + \frac{(\mathbf{-1})^4}{4}\right] + \left[\frac{\mathbf{1}^3}{3} - \frac{\mathbf{1}^4}{4}\right] - \left[\frac{\mathbf{0}^3}{3} - \frac{\mathbf{0}^4}{4}\right]$$

Which, when simplifying, yields:

$$V[X] = \left[\frac{1}{3} - \frac{1}{4}\right] + \left[\frac{1}{3} - \frac{1}{4}\right] = \left[\frac{2}{3} - \frac{2}{4}\right] = \frac{8}{12} - \frac{6}{12} = \frac{2}{12} = \frac{1}{6}$$

Answer: $\boldsymbol{E[X] = 0}$ **and** $\boldsymbol{V[X] = \frac{1}{6}}$

Question 4. Let p (which is a RV) be the price of a good such that $p \sim U[0, 10]$. The supply level s of the good is related to the price via,

$$\boldsymbol{s = 0.003p^2.}$$

Find the expected supply level both,

i) Directly

In order to find the expected supply level, we need to calculate the expected value of the variable s. Using the rules of expectations, we can write it as:

$$E[s] = E[0.003p^2] = 0.003E[p^2] = 0.003\int_0^{10} p^2 f(p)\, dp$$

We are told that variable p follows a uniform distribution between 0 and 10. The PDF of a uniform distribution is given by the following equation:

$$f(p) = \frac{1}{b-a}$$

Where b and a are, respectively, the highest and the lowest value that p can take. In our case, then, the PDF of p is given by the following equation:

$$f(p) = \frac{1}{10-0} = \frac{1}{10}$$

Substituting this value in the equation that describes the expected value of s, we get:

$$E[s] = 0.003\int_0^{10} p^2 \frac{1}{10} dp = \frac{0.003}{10}\int_0^{10} p^2\, dp$$

Solving the integral, we get:

$$E[s] = 0.0003\left[\frac{p^3}{3}\right]_0^{10}$$

Applying the First Fundamental Theorem of Calculus, we get:

$$E[s] = 0.0003\left[\frac{\mathbf{10}^3}{3} - \frac{\mathbf{0}^3}{3}\right]$$

Which, if simplifying, is equivalent to:

$$E[s] = 0.0003\frac{1000}{3} = \frac{0.3}{3} = 0.1$$

ii) By first determining the distribution of *s*

Theorem 3 in section 2.6 of the lecture notes states the following. Given that the PDF of X is $f(x)$, then the PDF of $Y = h(X)$ can be found by using the following formula:

$$g(y) = f(x) * \left|\frac{dx}{dy}\right|$$

Put simply, whenever we do not know the PDF of the variable Y but we do know that variable Y depends on X, we can work out the PDF of Y given the PDF of X and X itself.

In our case, we want to first determine the distribution of s and, afterwards, compute the PDF of s. Recalling that $s = 0.003p^2$, we can write p in terms of s:

$$p = \left(\frac{s}{0.003}\right)^{\frac{1}{2}}$$

As well, from the previous section we know that, given that p is a RV following a uniform distribution between 0 and 10, the PDF of p is given by:

$$f(p) = \frac{1}{10}$$

Given the two previous equations, we can rewrite the formula given by theorem 3 in terms of p and s:

$$g(s) = f(p) * \left|\frac{dp}{ds}\right|$$

Given that we have explicitly written p in terms of s before, we can directly compute $\frac{dp}{ds}$:

$$\frac{dp}{ds} = \frac{1}{2} * \left(\frac{s}{0.003}\right)^{-\frac{1}{2}} * \frac{1}{0.003}$$

Which, after simplifying, becomes:

$$\frac{dp}{ds} = \frac{1}{0.006} * \left(\frac{s}{0.003}\right)^{-\frac{1}{2}}$$

Noting that $x^{-a} = \frac{1}{x^a}$, we can rewrite the previous expression as:

$$\frac{dp}{ds} = \frac{1}{0.006} * \frac{1}{\left(\frac{s}{0.003}\right)^{\frac{1}{2}}} = \frac{1}{0.006} * \frac{1}{\frac{s^{\frac{1}{2}}}{0.003^{\frac{1}{2}}}} = \frac{0.003^{\frac{1}{2}}}{0.006} * \frac{1}{s^{\frac{1}{2}}} = \frac{0.003^{\frac{1}{2}}}{0.006} * s^{-\frac{1}{2}}$$

Substituting the previous result into $g(s)$, we get:

$$g(s) = \frac{1}{10}\left|\frac{0.003^{\frac{1}{2}}}{0.006} * s^{-\frac{1}{2}}\right|$$

As p has to lie within 0 and 10 and $p = \left(\frac{s}{0.003}\right)^{\frac{1}{2}} = \frac{\sqrt{s}}{0.003^{\frac{1}{2}}}$, then s is bounded to be positive. In practical terms, this implies that the absolute value will not change the function. Hence, we can rewrite $g(s)$ as:

$$g(s) = \frac{1}{10} * \frac{0.003^{\frac{1}{2}}}{0.006} * s^{-\frac{1}{2}} = \frac{0.003^{\frac{1}{2}}}{0.06} s^{-\frac{1}{2}}$$

Before start calculating the expected value of s, we need to show that $g(s)$ represents a valid PDF. To do that, the following rule has to hold true:

$$\int_0^x g(s)\, ds = \int_0^x \frac{0.003^{\frac{1}{2}}}{0.06} s^{-\frac{1}{2}} ds = 1$$

Which, operating, is shown to be equivalent to:

$$\int_0^x g(s)\, ds = \frac{0.003^{\frac{1}{2}}}{0.06} \int_0^x s^{-\frac{1}{2}} ds = 1$$

Solving the integral, we get:

$$\int_0^x g(s)\, ds = \frac{0.003^{\frac{1}{2}}}{0.06} \left[\frac{s^{\frac{1}{2}}}{\frac{1}{2}}\right]_0^x = 1$$

Applying the First Fundamental Theorem of Calculus, we get:

$$\int_0^x g(s)\,ds = \frac{0.003^{\frac{1}{2}}}{0.06}\left[\frac{\boldsymbol{x}^{\frac{1}{2}}}{\frac{1}{2}} - \frac{\boldsymbol{0}^{\frac{1}{2}}}{\frac{1}{2}}\right] = 1$$

Which, when simplifying, is equivalent to:

$$\int_0^x g(s)\,ds = \frac{0.003^{\frac{1}{2}}}{0.06}\frac{x^{\frac{1}{2}}}{\frac{1}{2}} = 1$$

Isolating x, we get:

$$x^{\frac{1}{2}} = \frac{\frac{1}{2} * 0.06}{0.003^{\frac{1}{2}}}$$

Squaring both sides yields:

$$x = \frac{0.06^2}{2^2 * 0.003} = \frac{0.0036}{0.012} = 0.3$$

Hence, for $g(s)$to be a valid PDF, then $0 \leq s \leq 0.3$ has to hold true. Knowing this fact, we can now write the expected value of s as:

$$E[s] = \int_0^{0.3} sg(s)\,ds = \int_0^{0.3} s\frac{0.003^{\frac{1}{2}}}{0.06}s^{-\frac{1}{2}}\,ds = \int_0^{0.3}\frac{0.003^{\frac{1}{2}}}{0.06}s^{\frac{1}{2}}\,ds = \frac{0.003^{\frac{1}{2}}}{0.06}\int_0^{0.3} s^{\frac{1}{2}}\,ds$$

Solving the integral, we get:

$$E[s] = \frac{0.003^{\frac{1}{2}}}{0.06}\left[\frac{s^{\frac{3}{2}}}{\frac{3}{2}}\right]_{\mathbf{0}}^{\mathbf{0.3}}$$

Applying the First fundamental Theorem of Calculus, we get:

$$E[s] = \frac{0.003^{\frac{1}{2}}}{0.06}\left[\frac{\mathbf{0.3}^{\frac{3}{2}}}{\frac{3}{2}} - \frac{\mathbf{0}^{\frac{3}{2}}}{\frac{3}{2}}\right] = \frac{0.003^{\frac{1}{2}}}{0.06}\frac{0.3^{\frac{3}{2}}}{\frac{3}{2}} = \frac{2}{3} * \frac{0.003^{\frac{1}{2}}}{0.006} * 0.3^{\frac{3}{2}} = 0.1$$

Answer: $\boldsymbol{E[s] = 0.1}$

L11MEE TUTORIALS – SEMESTER 2 – TUTORIAL 3

By ERNESTO M. GAVASSA PEREZ*

---

*This document was compiled based on the answers provided by Professor Patrick Marsh. Hence, all the credit should be given to him

Question 1. Let $z_1, z_2, \dots, z_n$ be a sequence of independent normal (IN) RVs having means μ_i and variances σ_i^2, for $i = 1, \dots, n$. Deduce that the RV

$$\boldsymbol{Q = \sum_{i=1}^{n} \frac{(z_i - \mu_i)^2}{\sigma_i^2}}$$

Has a $\chi^2_{(n)}$ distribution.

By the definition of a Chi-Square Distribution of section 3.2 in the lecture notes, we know that the sum of several independent standardised squared normal RVs follows a Chi-Square Distribution with n degrees of freedom. More formally, if $X_i \sim N(0,1)$ for $i = 1, \dots, n$, then it follows that:

$$Q \sim X^2_{(n)}, where\ Q = \sum_{i=1}^{n} X_i^2$$

In order to deduce that the RV Q has a $X^2_{(n)}$ distribution, we need to demonstrate that each element within the summation is just the square of a Normal Distribution with mean 0 and variance1. Namely, that each RV of the type X_i follows a standardised Normal Distribution. The first step is to rewrite the summation in the following way:

$$Q = \sum_{i=1}^{n} \frac{(z_i - \mu_i)^2}{\sigma_i^2} = \sum_{i=1}^{n} \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2$$

The second step is to note that each element within the summation is just a transformation of the z_i random variables described in the questions. To make the notation consistent, let us name the transformed RV in the following way:

$$X_i = \frac{z_i - \mu_i}{\sigma_i}$$

Hence, we can rewrite Q by substituting each element within the summation by X_i:

$$Q = \sum_{i=1}^{n} \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 = \sum_{i=1}^{n} (X_i)^2$$

Now, it should be easier to see that in order for the definition of the section 3.2 to apply, we need to demonstrate that the RVs X_i follow a normal distribution with mean 0 and variance 1. In order to do that, we use the rules of expectations and the knowledge that $E[z_i] = \mu_i$ and $V[z_i] = \sigma_i^2$.

First, let us write the expected value of X_i as:

$$E[X_i] = E\left[\frac{z_i - \mu_i}{\sigma_i}\right]$$

We can divide what lies within the expectation into two elements:

$$E[X_i] = E\left[\frac{z_i - \mu_i}{\sigma_i}\right] = E\left[\frac{z_i}{\sigma_i} - \frac{\mu_i}{\sigma_i}\right]$$

By the rules of expectations, we know that the expected value of a constant is the constant itself ($E[b] = b$) and that the expected value of a RV times a constant is the constant times the expected value of the RV ($E[aX_i] = aE[X_i]$). Together, these two rules imply:

$$E[X_i] = E\left[\frac{z_i}{\sigma_i} - \frac{\mu_i}{\sigma_i}\right] = \frac{1}{\sigma_i}E[z_i] - \frac{\mu_i}{\sigma_i}$$

Where $\frac{1}{\sigma_i} = a$ and $\frac{\mu_i}{\sigma_i} = b$, using the notation of the aforementioned rules. As the question states that the expected value of the RV $E[z_i] = \mu_i$, we can substitute that in the previous equation to get:

$$E[X_i] = \frac{\mu_i}{\sigma_i} - \frac{\mu_i}{\sigma_i} = 0$$

Hence, the first condition for the definition in section 3.2 to apply is satisfied. Now, let us write the variance of X_i as:

$$V[X_i] = E[(X_i - E[X_i])^2]$$

As we have just previously found that $E[X_i] = 0$, we can substitute that in the previous equation to get:

$$V[X_i] = E[(X_i - 0)^2] = E[X_i^2]$$

Hence, we need to demonstrate now that $V[X_i] = 1$. For that, let us rewrite the previous expression as:

$$V[X_i] = E[X_i^2] = E\left[\left(\frac{z_i - \mu_i}{\sigma_i}\right)^2\right] = E\left[\left(\frac{z_i}{\sigma_i} - \frac{\mu_i}{\sigma_i}\right)^2\right]$$

Expanding the expression within the brackets, we get:

$$V[X_i] = E\left[\left(\frac{z_i}{\sigma_i} - \frac{\mu_i}{\sigma_i}\right)^2\right] = E\left[\left(\frac{z_i}{\sigma_i}\right)^2 + \left(\frac{\mu_i}{\sigma_i}\right)^2 - 2z_i\mu_i\right] = E\left[\frac{z_i^2}{\sigma_i^2} + \frac{\mu_i^2}{\sigma_i^2} - 2\frac{z_i}{\sigma_i}\frac{\mu_i}{\sigma_i}\right]$$

As $\frac{1}{\sigma_i^2}$, $2\frac{\mu_i}{\sigma_i^2}$ and $\frac{\mu_i^2}{\sigma_i^2}$ are constants, and applying the rules mentioned before, the previous expression is equivalent to:

$$V[X_i] = E\left[\frac{z_i^2}{\sigma_i^2} + \frac{\mu_i^2}{\sigma_i^2} - 2\frac{z_i\mu_i}{\sigma_i^2}\right] = \frac{1}{\sigma_i^2}E[z_i^2] + \frac{\mu_i^2}{\sigma_i^2} - 2\frac{\mu_i}{\sigma_i^2}E[z_i]$$

As $E[z_i] = \mu_i$, the previous equation becomes:

$$V[X_i] = \frac{1}{\sigma_i^2}E[z_i^2] + \frac{\mu_i^2}{\sigma_i^2} - 2\frac{\mu_i}{\sigma_i^2}\mu_i = \frac{1}{\sigma_i^2}E[z_i^2] + \frac{\mu_i^2}{\sigma_i^2} - 2\frac{\mu_i^2}{\sigma_i^2}$$

Simplifying, we get:

$$V[X_i] = \frac{1}{\sigma_i^2}E[z_i^2] - \frac{\mu_i^2}{\sigma_i^2}$$

At the moment, $E[z_i^2]$ is unknown to us. However, recall that we know that $z_i \sim N(\mu_i, \sigma_i^2)$. Hence, it must be that:

$$V[z_i] = \sigma_i^2$$

Substituting $V[z_i]$ by the generic formula, we get:

$$V[z_i] = E[(z_i - E[z_i])^2] = \sigma_i^2$$

As the question states that $E[z_i] = \mu_i$, when substituting such equivalence in the previous equation we get:

$$V[z_i] = E[(z_i - \mu_i)^2] = \sigma_i^2$$

Expanding the parenthesis, we get:

$$V[z_i] = E[(z_i - \mu_i)^2] = E[z_i^2 + \mu_i^2 - 2z_i\mu_i] = \sigma_i^2$$

By noticing that μ_i^2 and $2\mu_i$ are constants, and applying the aforementioned rules of expectations, we get:

$$V[z_i] = E[z_i^2 + \mu_i^2 - 2z_i\mu_i] = E[z_i^2] + \mu_i^2 - 2\mu_i E[z_i] = \sigma_i^2$$

By substituting $E[z_i] = \mu_i$ – which the question states – in the previous equation, we get:

$$V[z_i] = E[z_i^2] + \mu_i^2 - 2\mu_i E[z_i] = E[z_i^2] + \mu_i^2 - 2\mu_i\mu_i = \sigma_i^2$$

Simplifying, we get:

$$V[z_i] = E[z_i^2] + \mu_i^2 - 2\mu_i\mu_i = E[z_i^2] + \mu_i^2 - 2\mu_i^2 = \sigma_i^2$$

Which is equivalent to:

$$V[z_i] = E[z_i^2] + \mu_i^2 - 2\mu_i^2 = E[z_i^2] - \mu_i^2 = \sigma_i^2$$

By isolating $E[z_i^2]$, which was the unknown of the simplified expression of $V[X_i]$, we get:

$$E[z_i^2] = \sigma_i^2 + \mu_i^2$$

Recall the formula we got for $V[X_i]$ was:

$$V[X_i] = \frac{1}{\sigma_i^2} E[z_i^2] - \frac{\mu_i^2}{\sigma_i^2}$$

Substituting the formula for $E[z_i^2]$ in $V[X_i]$, we get:

$$V[X_i] = \frac{1}{\sigma_i^2} E[z_i^2] - \frac{\mu_i^2}{\sigma_i^2} = \frac{1}{\sigma_i^2}(\sigma_i^2 + \mu_i^2) - \frac{\mu_i^2}{\sigma_i^2}$$

By expanding the first term, we get:

$$V[X_i] = \frac{1}{\sigma_i^2}\sigma_i^2 + \frac{1}{\sigma_i^2}\mu_i^2 - \frac{\mu_i^2}{\sigma_i^2}$$

Which is equivalent to:

$$V[X_i] = \frac{\sigma_i^2}{\sigma_i^2} + \frac{\mu_i^2}{\sigma_i^2} - \frac{\mu_i^2}{\sigma_i^2}$$

As $\frac{\sigma_i^2}{\sigma_i^2} = 1$ and $\frac{\mu_i^2}{\sigma_i^2} - \frac{\mu_i^2}{\sigma_i^2} = 0$, the previous expression collapses to:

$$V[X_i] = 1$$

Hence, we have demonstrated that, given that the RV $z_i \sim N(\mu_i, \sigma_i^2)$, then the standardised RV $X_i = \frac{z_i - \mu_i}{\sigma_i} \sim N(0,1)$. Hence, the definition in section 3.2 of the lecture note applies and the RV Q as defined in the question follows a Chi-Square Distribution with n degrees of freedom; where n is the number of standardised RVs of the type z_i.

Question 2. If c is a constant then demonstrate that,

i) $\boldsymbol{V}[\boldsymbol{X}+\boldsymbol{c}]=\boldsymbol{V}[\boldsymbol{X}]$

The generic formula for$E[X]$ is:

$$E[X]=\int_{-\infty}^{\infty} xf(x)dx$$

Hence, we can write the expected value of $X+c$ as:

$$E[X+c]=\int_{-\infty}^{\infty}(x+c)f(x)dx$$

Expanding the parenthesis, we get:

$$E[X+c]=\int_{-\infty}^{\infty}[xf(x)+cf(x)]dx$$

Using the rules of calculus (more specifically, the fact that $\int \big(f(x)+g(x)\big)dx=\int f(x)dx+\int g(x)dx$), we get:

$$E[X+c]=\int_{-\infty}^{\infty} xf(x)dx+\int_{-\infty}^{\infty} cf(x)dx$$

As c is just a constant, we can take it out of the second integral to get:

$$E[X+c]=\int_{-\infty}^{\infty} xf(x)dx+c\int_{-\infty}^{\infty} f(x)dx$$

As $\int_{-\infty}^{\infty} f(x)dx=1$ for $f(x)$ to be a valid PDF, we can substitute the second integral by 1 in order to get:

$$E[X+c]=\int_{-\infty}^{\infty} xf(x)dx+c$$

Substituting $E[X]=\int_{-\infty}^{\infty} xf(x)dx$ in the previous equation, we get:

$$E[X+c] = E[X] + c$$

Now, we can make use of the following generic formula:

$$V[X] = E[(X - E[X])^2]$$

And substitute X by $X + c$ and $E[X + c]$ by the expression we get above to get:

$$V[X+c] = E[(X - E[X+c])^2] = E\left[\left(X + c - (E[X] + c)\right)^2\right]$$

Expanding the expression within the parenthesis, we get:

$$V[X+c] = E[(X + c - E[X] - c)^2]$$

The c's cancel out, yielding the following expression:

$$V[X+c] = E[(X - E[X])^2] = V[X]$$

Hence, as shown, the variance of the RVs X and $X + c$ are identical.

ii) $\boldsymbol{V[cX] = c^2 V[X]}$

Using the general formula for the variance of a RV, we can write the variance of cX as:

$$V[cX] = E[(cX - E[cX])^2]$$

Using the formula for the expectation of an RV, we can write the expectation of the RV cX as:

$$E[cX] = \int_{-\infty}^{\infty} cXf(X)dx$$

As c is a constant, we can take it out of the integral to get:

$$E[cX] = c\int_{-\infty}^{\infty} Xf(X)dx = cE[X]$$

Substituting for the expression of $E[cX]$ found above, we can rewrite the variance of cX as:

$$V[cX] = E[(cX - E[cX])^2] = E[(cX - cE[X])^2]$$

By using c as a common factor, we can rewrite the previous expression as:

$$V[cX] = E[(cX - E[cX])^2] = E\left[\left(c(X - E[X])\right)^2\right]$$

By the rules of calculus, we know that $\left(cf(x)\right)^2 = c^2 f(x)^2$. Hence, we can rewrite the previous expression as:

$$V[cX] = E[(cX - E[cX])^2] = E[c^2(X - E[X])^2]$$

Because of the expression we found above for $E[cX]$, we can deduce that, as c^2 is a constant, then it can be taken out of the expectation:

$$V[cX] = E[c^2(X - E[X])^2] = c^2 E[(X - E[X])^2] = c^2 V[X]$$

It, then, follows that $V[cX] = c^2 V[X]$.

Question 3. Show that if X is a RV having a PDF which is symmetric around $x = 0$ then $E[X] = 0$. Extend this analysis to the case where $f(x)$ is symmetric around $x = a$ to show then that $E[X] = a$.

A PDF of the RV X is defined generically as $f(x)$. It follows, then, that for the RV X to be symmetric then the equivalence $f(x) = f(-x)$ has to hold true. As we want to calculate the expectation of the RV X, let us write the generic formula below:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Using the rules of calculus, we can rewrite it as:

$$E[X] = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{\infty} xf(x)dx$$

As we do not know the functional form of $f(x)$, we cannot solve the integrals above. However, to make the expectation 0 it will suffice to show that, for any generic function $f(x)$, the first and the second integral are equal in magnitude but opposite in sign. To see this, we can apply the rule stating that $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$[2] to the second integral to get:

$$E[X] = \int_{-\infty}^{0} xf(x)dx - \int_{\infty}^{0} xf(x)dx$$

Now, we subsequently apply the rule stating that $\int_a^b f(x)\,dx = \int_{-a}^{-b} f(-x)(-\,dx)$[3] to the second integral to get:

[2] To see this, and assuming that $F(x)$ is the generic solution to the proposed integral, we can apply the First Fundamental Theorem of Calculus to both sides to get the following expression:

$$F(b) - F(a) = -[F(a) - F(b)]$$

Which, after expanding the brackets of the right hand side, becomes:

$$F(b) - F(a) = -F(a) - \big(-F(b)\big) = -F(a) + F(b)$$

[3] Following the previous footnote, and applying the first Fundamental Theorem of Calculus to each of the hand sides of the stated rule, we get:

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

$$\int_{-a}^{-b} f(-x)\,(-dx) = F\big(-(-b)\big) - F\big(-(-a)\big) = F(b) - F(a)$$

$$E[X] = \int_{-\infty}^{0} xf(x)dx - \int_{-\infty}^{0} (-x)f(-x)(-dx)$$

As $f(x) = f(-x)$ by symmetry, and $(-x)(-dx) = (-1)(x)(-1)(dx) = xdx$, we can rewrite the previous expression as:

$$E[X] = \int_{-\infty}^{0} xf(x)dx - \int_{-\infty}^{0} (x)f(x)(dx) = 0$$

In order to generalise the results to a function symmetric around a, we first need to note that $f(x-a) = f\big(-(x-a)\big)$. By imposing a new RV Z to be defined as $Z = X - a$, and applying the rules of expectations, we know that:

$$E[Z] = E[X-a] = E[X] - a$$

Since now the function is symmetric around $z = x - a$ and not around x, this implies that, now, $f(x-a) = f(z) = f\big(-(x-a)\big) = f(-z)$. Hence, the argument constructed before can now apply for z:

$$E[Z] = \int_{-\infty}^{0} zf(z)dz + \int_{0}^{\infty} zf(z)dz$$

Which, applying all the rules mentioned before, becomes:

$$E[Z] = \int_{-\infty}^{0} zf(z)dz - \int_{-\infty}^{0} (z)f(z)(dz) = 0$$

Substituting this in the expression for the expectation of the RV Z, we get:

$$E[Z] = 0 = E[X] - a$$

By isolating $E[X]$, we finally get:

$$0 + a = E[X] = a$$

Question 4. Show that

$$\sum_{i=1}^{n}(\boldsymbol{x_i} - \overline{\boldsymbol{x}})^2 = \sum_{i=1}^{n} \boldsymbol{x_i^2} - \boldsymbol{n\overline{x}^2},$$

Where $\overline{\boldsymbol{x}} = \frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{i=1}^{n} \boldsymbol{x_i}$.

To show that the stated equivalence holds, we first expand the left hand side of the equation to get:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \bar{x}^2 - 2 * \sum_{i=1}^{n} x_i\bar{x}$$

As $\bar{x}$ is a constant, we can bring it out of the summation in the third element of the previous equation to get[4]:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \bar{x}^2 - 2\bar{x} * \sum_{i=1}^{n} x_i$$

Additionally, $\bar{x}^2$ is a constant as well. Hence, summing $\bar{x}^2$ n times yields $\bar{x}^2$[5] :

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x} * \sum_{i=1}^{n} x_i$$

Using the formula for $\bar{x}$ provided in the question ($\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$), we can isolate $\sum_{i=1}^{n} x_i$ to get:

$$\sum\nolimits_{i=1}^{n} x_i = n * \bar{x}$$

Substituting this expression in the previous formula, we get:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x} * n * \bar{x}$$

---

[4] To see this, note that $\sum_{i=1}^{n} x_i\bar{x}$ can be rewritten as $x_1\bar{x} + x_2\bar{x} + \cdots + x_n\bar{x}$. Taking $\bar{x}$ as a common factor, we get: $\bar{x}(x_1 + x_2 + \cdots + x_n) = \bar{x}\sum_{i=1}^{n} x_i$

[5] To see this, note that $\sum_{i=1}^{n} \bar{x}^2$ can be rewritten as $\bar{x}^2 + \bar{x}^2 + \cdots + \bar{x}^2$. Using $\bar{x}^2$ as a common factor, this expression is equivalent to $\bar{x}^2 * (1 + 1 + \cdots + 1)$. Noticing that $\bar{x}^2$ was added n times, this is equivalent to $n\bar{x}^2$.

Which simplifies to:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2n * \bar{x}^2 = \sum_{i=1}^{n} x_i^2 - n * \bar{x}^2$$

Question 5. A RV X has an exponential function if

$$\boldsymbol{f(x) = \frac{e^{-\frac{x}{\beta}}}{\beta}, \qquad x > 0.}$$

Determine the cumulative distribution function of X and also find both $E[X]$ and $V[X]$. Suppose $\beta = 100$, find the value c such that

$$\mathbf{Pr}[\boldsymbol{X \geq c}] = \frac{\mathbf{1}}{\mathbf{2}}.$$

Determine the cumulative distribution function of X

The Cumulative Distribution function (CDF for short) of a RV X is to be defined as:

$$Pr[X \leq z] = \int_{-\infty}^{z} f(x)dx$$

Given that the lower bound of the RV X is 0 and the $f(x)$ described in the question, the CDF of X is:

$$Pr[X \leq z] = \int_{0}^{z} \frac{e^{-\frac{x}{\beta}}}{\beta} dx$$

We can also write it as follows

$$Pr[X \leq z] = \int_{0}^{z} \left(\frac{1}{\beta} e^{-\frac{x}{\beta}}\right) dx$$

We can substitute $1 = (-1) * (-1)$ and take one of the -1's outside of the integral to get:

$$Pr[X \leq z] = (-1) \int_{0}^{z} \left((-1)\frac{1}{\beta} e^{-\frac{x}{\beta}}\right) dx = -\int_{0}^{z} \left(-\frac{1}{\beta} e^{-\frac{x}{\beta}}\right) dx$$

By noticing that $-\frac{1}{\beta} = \frac{d\left(-\frac{x}{\beta}\right)}{dx}$; and, hence, noticing that $-\frac{1}{\beta} e^{-\frac{x}{\beta}} = \frac{d\left(e^{-\frac{x}{\beta}}\right)}{dx}$, the previous expression is equivalent to:

$$Pr[X \leq z] = -\int_{0}^{z} \left(\frac{d\left(e^{-\frac{x}{\beta}}\right)}{dx} \right) dx$$

Hence, using the rules of calculus (more specifically, the fact that $\int f(x)dx = \int \frac{d(F(x))}{dx} dx = F(x) + C$, the CDF is defined by the expression below:

$$Pr[X \leq z] = -\left[e^{-\frac{x}{\beta}}\right]_{0}^{z}$$

Using the First Fundamental Theorem of Calculus, we get:

$$Pr[X \leq z] = -\left\{\left[e^{-\frac{z}{\beta}}\right] - \left[e^{-\frac{0}{\beta}}\right]\right\} = e^{-\frac{0}{\beta}} - e^{-\frac{z}{\beta}}$$

Using the rules of calculus (more specifically, noting that $x^{-a} = \frac{1}{x^{a}}$), the previous expression is equivalent to:

$$Pr[X \leq z] = \frac{1}{e^{\frac{0}{\beta}}} - e^{-\frac{z}{\beta}} = \frac{1}{e^{0}} - e^{-\frac{z}{\beta}} = 1 - e^{-\frac{z}{\beta}}$$

Which describes the CDF of the RV X.

Determine the $\boldsymbol{E[X]}$

Using the formula for the expectation of a random variable, we can write the expectation of X as:

$$E[X] = \int_0^\infty x \frac{e^{-\frac{x}{\beta}}}{\beta} dx$$

This integral needs to be solved using the technique known as integration by parts. Recalling the formula for integrating by parts

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

And setting $f(x) = x$ and $g'(x) = \frac{e^{-\frac{x}{\beta}}}{\beta}$, the two remaining elements appearing in the previous expression are:

$$f'(x) = \frac{df(x)}{dx} = \frac{dx}{dx} = 1$$

$$g(x) = \int g'(x)dx = \int \frac{e^{-\frac{x}{\beta}}}{\beta} dx = -\int -\frac{1}{\beta} e^{-\frac{x}{\beta}} dx = -\int \frac{d\left(e^{-\frac{x}{\beta}}\right)}{dx} dx = -e^{-\frac{x}{\beta}}$$

Substituting $f(x)$, $f'(x)$, $g(x)$ and $g'(x)$, we get:

$$E[X] = \int_0^\infty \left(x \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx = \left[-xe^{-\frac{x}{\beta}}\right]_0^\infty - \int_0^\infty \left(1 * \left(-e^{-\frac{x}{\beta}}\right)\right) dx$$

Now, substituting $1 = \frac{\beta}{\beta}$ within the integral in the right hand side, we get:

$$E[X] = \int_0^\infty \left(x \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx = \left[-xe^{-\frac{x}{\beta}}\right]_0^\infty - \int_0^\infty \left(-\frac{\beta}{\beta} * e^{-\frac{x}{\beta}}\right) dx$$

As both the numerator and denominator of $\frac{\beta}{\beta}$ are constants, we can take whichever of them outside the integral. Hence, by taking the numerator outside of the integral of the right hand side, we get:

$$E[X]=\int_0^\infty\left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx=\left[-xe^{-\frac{x}{\beta}}\right]_0^\infty-\beta\int_0^\infty\left(-\frac{1}{\beta}*e^{-\frac{x}{\beta}}\right)dx$$

By noticing, as before, that $-\frac{1}{\beta}*e^{-\frac{x}{\beta}}=\frac{d\left(e^{-\frac{x}{\beta}}\right)}{dx}$, the solution to the integral of the right hand side can be calculated accordingly:

$$E[X]=\int_0^\infty\left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx=\left[-xe^{-\frac{x}{\beta}}\right]_0^\infty-\beta\left[e^{-\frac{x}{\beta}}\right]_0^\infty$$

Applying the First Fundamental Theorem of Calculus, we get:

$$E[X]=\int_0^\infty\left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx=-\infty e^{-\frac{\infty}{\beta}}+0e^{-\frac{0}{\beta}}-\beta\left[e^{-\frac{\infty}{\beta}}-e^{-\frac{0}{\beta}}\right]$$

Using the rules of calculus (more specifically, using the fact that $x^{-a}=\frac{1}{x^a}$), we can rewrite the previous expression as:

$$E[X]=\int_0^\infty\left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx=-\frac{\infty}{e^{\frac{\infty}{\beta}}}+\frac{0}{e^{\frac{0}{\beta}}}-\beta\left[\frac{1}{e^{\frac{\infty}{\beta}}}-\frac{1}{e^{\frac{0}{\beta}}}\right]=-\frac{\infty}{e^\infty}+\frac{0}{e^0}-\beta\left[\frac{1}{e^\infty}-\frac{1}{e^0}\right]$$

By noticing that $e^\infty=\infty$ and $e^0=1$, we can rewrite the previous equation as:

$$E[X]=\int_0^\infty\left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx=-\frac{\infty}{\infty}+\frac{0}{1}-\beta\left[\frac{1}{\infty}-\frac{1}{1}\right]$$

Even if it is tempting to assume that $\frac{\infty}{\infty}=1$, $\frac{\infty}{\infty}$ it is an **indeterminate form**. In order to know the real value of $\frac{\infty}{\infty}$ in our setting, **we need to apply l'Hôpital's Rule**, which states the following:

If f and g are differentiable functions for all values of x larger than a given number, a, and it is the case that both $\lim_{x\to a}f(x)=\infty$ and $\lim_{x\to a}g(x)=\infty$, then the following equivalence holds true:

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}$$

In our case, $\lim_{x\to\infty} x = \infty$ and $\lim_{x\to\infty} e^{\frac{x}{\beta}} = \infty$. Hence, from l'Hôpital's Rule it follows that:

$$\lim_{x\to\infty} \frac{x}{e^{\frac{x}{\beta}}} = \lim_{x\to\infty} \frac{1}{\frac{1}{\beta}e^{\frac{x}{\beta}}} = \frac{1}{\frac{1}{\beta}e^{\frac{\infty}{\beta}}} = \frac{1}{\frac{1}{\beta}e^{\infty}} = \frac{1}{\frac{1}{\beta}\infty} = \frac{1}{\infty} = 0$$

Hence, we can rewrite the expectation of X as:

$$E[X] = \int_0^\infty \left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx = -\frac{\infty}{\infty} + \frac{0}{1} + \beta\left[\frac{1}{\infty} - \frac{1}{1}\right] = -0 + 0 - \beta[0-1]$$

Which, when simplifying, is equivalent to:

$$E[X] = \int_0^\infty \left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx = \beta$$

Determine $\boldsymbol{V}[\boldsymbol{X}]$

In order to calculate the variance of the RV X, we use the general formula:

$$V[X] = E[(X - E[X])^2]$$

As we already know that $E[X] = \beta$, we can substitute that quantity in the previous expression to get:

$$V[X] = E[(X - \beta)^2] = \int_0^\infty \left[(x-\beta)^2 * \frac{e^{-\frac{x}{\beta}}}{\beta}\right] dx$$

Expanding the parenthesis, we get:

$$V[X] = \int_0^\infty \left[(x^2 + \beta^2 - 2\beta x) * \frac{e^{-\frac{x}{\beta}}}{\beta}\right] dx$$

Applying the rules of calculus (more specifically, the fact that $\int \big(f(x) + g(x)\big)dx = \int f(x)dx + \int g(x)dx$), we can rewrite the previous equation as three different integrals:

$$V[X] = \int_0^\infty \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx + \int_0^\infty \left(\beta^2 * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx - \int_0^\infty \left(2\beta x * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx$$

As β^2 and 2β are just constants within the second and third integrals respectively, we can take them out to get:

$$V[X] = \int_0^\infty \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx + \beta^2 \int_0^\infty \left(\frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx - 2\beta \int_0^\infty \left(x * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx$$

As $E[X] = \int_0^\infty \left(x \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx = \beta$, we can substitute the third integral by β to get:

$$V[X] = \int_0^\infty \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx + \beta^2 \int_0^\infty \left(\frac{e^{-\frac{x}{\beta}}}{\beta}\right) dx - 2\beta\beta$$

Which is equivalent to:

$$V[X] = \int_0^{\infty} \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx + \beta^2 \int_0^{\infty} \left(\frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx - 2\beta^2$$

Also, in order for $f(x) = \frac{e^{-\frac{x}{\beta}}}{\beta}$ to be a valid PDF, then it follows that $\int_0^{\infty} \left(\frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx = 1$. Hence, substituting the second integral by 1, we get:

$$V[X] = \int_0^{\infty} \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx + \beta^2 * (1) - 2\beta^2$$

Which is identical to:

$$V[X] = \int_0^{\infty} \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx + \beta^2 - 2\beta^2$$

Simplifying, we get:

$$V[X] = \int_0^{\infty} \left(x^2 * \frac{e^{-\frac{x}{\beta}}}{\beta} \right) dx - \beta^2$$

Again, the remaining integral needs to be solved with the *integration by parts* technique. Fixing $f(x) = x^2$ and $g'(x) = \frac{e^{-\frac{x}{\beta}}}{\beta}$, the two remaining elements to find the relevant integral are:

$$f'(x) = 2x$$

$$g(x) = \int g'(x)dx = -e^{-\frac{x}{\beta}}$$

Hence, we can rewrite the variance of the RV X as:

$$V[X] = \left[-x^2 e^{-\frac{x}{\beta}} \right]_0^{\infty} - \int_0^{\infty} \left(-2x * e^{-\frac{x}{\beta}} \right) dx - \beta^2$$

As -2 is a constant, we can take it out of the integral to get:

$$V[X] = \left[-x^2 e^{-\frac{x}{\beta}} \right]_0^{\infty} + 2 \int_0^{\infty} \left(x e^{-\frac{x}{\beta}} \right) dx - \beta^2$$

Knowing that $\int_0^\infty \left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx = \beta$, and noting that the $\frac{1}{\beta}$ within the integral is a constant, we can deduce that:

$$\int_0^\infty \left(x\frac{e^{-\frac{x}{\beta}}}{\beta}\right)dx = \frac{1}{\beta}\int_0^\infty \left(xe^{-\frac{x}{\beta}}\right)dx = \beta$$

Isolating the integral, we get:

$$\int_0^\infty \left(xe^{-\frac{x}{\beta}}\right)dx = [\beta\beta]_0^\infty = [\beta^2]_0^\infty$$

Substituting it into the variance of x, we get:

$$V[X] = \left[-x^2e^{-\frac{x}{\beta}}\right]_0^\infty + 2[\beta^2]_0^\infty - \beta^2$$

Applying the First Fundamental Theorem of Calculus, we get:

$$V[X] = -(\infty)^2e^{-\frac{\infty}{\beta}} + (0)^2e^{-\frac{0}{\beta}} + 2\beta^2 - \beta^2$$

Which, after simplifying, becomes:

$$V[X] = -\frac{\infty}{e^{\frac{\infty}{\beta}}} + \beta^2 = -\frac{\infty}{\infty} + \beta^2$$

Again, applying l'Hôpital's Rule for assessing the term with an indeterminate form, we get:

$$\lim_{x\to\infty}\frac{x^2}{e^{\frac{x}{\beta}}} = \lim_{x\to\infty}\frac{2x}{\frac{1}{\beta}e^{\frac{x}{\beta}}} = \frac{2*\infty}{\frac{1}{\beta}e^{\frac{\infty}{\beta}}} = \frac{\infty}{\frac{1}{\beta}e^{\infty}} = \frac{\infty}{\frac{1}{\beta}\infty} = \frac{\infty}{\infty}$$

By applying the l'Hôpital's Rule a second time, we get:

$$\lim_{x\to\infty}\frac{2x}{\frac{1}{\beta}e^{\frac{x}{\beta}}} = \lim_{x\to\infty}\frac{2}{\left(\frac{1}{\beta}\right)^2e^{\frac{x}{\beta}}} = \frac{2}{\left(\frac{1}{\beta}\right)^2e^{\frac{\infty}{\beta}}} = \frac{2}{\left(\frac{1}{\beta}\right)^2e^{\infty}} = \frac{2}{\left(\frac{1}{\beta}\right)^2\infty} = \frac{2}{\infty} = 0$$

Hence, the variance of the RV X is defined by the expression below:

$$V[X] = -\frac{\infty}{\infty} + \beta^2 = 0 + \beta^2 = \beta^2$$

Suppose $\boldsymbol{\beta = 100}$. Determine the value $\boldsymbol{c}$ such that

$$\mathbf{Pr}[\boldsymbol{X \geq c}] = \frac{\mathbf{1}}{\mathbf{2}}.$$

Finally, we are asked to find the value of c for which $\Pr[X \geq c] = \frac{1}{2}$, given that $\beta = 100$. First, let us rewrite $f(x)$ for the proposed value of β:

$$f(x) = \frac{e^{-\frac{x}{100}}}{100}$$

In order to find the proposed probability, we write the probability as a definite integral:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow \int_{c}^{\infty} \left(\frac{e^{-\frac{x}{100}}}{100} \right) dx = \frac{1}{2}$$

Multiplying the integrand by $1 = (1) * (-1)$, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow \int_{c}^{\infty} \left[(-1) * (-1) * \left(\frac{e^{-\frac{x}{100}}}{100} \right) \right] dx = \frac{1}{2}$$

Taking one of the -1’s out of the integral, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow -\int_{c}^{\infty} \left(-\frac{e^{-\frac{x}{100}}}{100} \right) dx = -\int_{c}^{\infty} \left(-\frac{1}{100} e^{-\frac{x}{100}} \right) dx = \frac{1}{2}$$

By noticing that $-\frac{1}{100} e^{-\frac{x}{100}} = \frac{d\left(e^{-\frac{x}{100}}\right)}{dx}$, and using the aforementioned rules of calculus, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow -\left[e^{-\frac{x}{100}} \right]_{c}^{\infty} = \frac{1}{2}$$

By applying the First Fundamental Theorem of Calculus, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow -\left[e^{-\frac{\infty}{100}} - e^{-\frac{c}{100}} \right] = \frac{1}{2}$$

Which, by noticing that $x^{-a} = \frac{1}{x^{a}}$, becomes:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow -\left[\frac{1}{e^{\frac{\infty}{100}}} - \frac{1}{e^{\frac{c}{100}}}\right] = -\left[\frac{1}{e^{\infty}} - \frac{1}{e^{\frac{c}{100}}}\right] = -\left[\frac{1}{\infty} - \frac{1}{e^{\frac{c}{100}}}\right] = \frac{1}{e^{\frac{c}{100}}} = \frac{1}{2}$$

Rearranging, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow 2 = e^{\frac{c}{100}}$$

Taking logs in both hand sides, we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow \ln(2) = \ln\left(e^{\frac{c}{100}}\right)$$

Which, by applying the rules of calculus, is equivalent to:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow \ln(2) = \frac{c}{100}\ln(e) = \frac{c}{100}$$

Hence, isolating c we get:

$$\Pr[X \geq c] = \frac{1}{2} \Longrightarrow 100\ln(2) = c = 69.315$$

L11MEE TUTORIALS – SEMESTER 2 – TUTORIAL 4

By ERNESTO M. GAVASSA PEREZ*

*This document was compiled based on the answers provided by Professor Patrick Marsh. Hence, all the credit should be given to him

1. A classical regression model is written,

$$\boldsymbol{y_t = \alpha + \beta x_t + \varepsilon_t, \quad \varepsilon_t \sim IIDN(0, \sigma^2), \quad t = 1, 2, \dots, n}$$

Where the values of $\boldsymbol{x_1, x_2, \dots, x_n}$ are not random.

(a) Derive the Ordinary Least Squares estimators of both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, say $\boldsymbol{\widehat{\alpha}}$ and $\boldsymbol{\widehat{\beta}}$.

By definition, we know that $y_t = \hat{y}_t + \hat{\varepsilon}_t$. Hence, we can isolate the residuals to get:

$$\hat{\varepsilon}_t = y_t - \hat{y}_t$$

More crucially, this equation just outlines the residual of each observation in our dataset; where each observational unit is described by the subscript t. In order to sum the residuals of all the observations in our data, we need to sum the residuals of each of the observations in our dataset, which we can write compactly as:

$$\sum_{t=1}^{n} \hat{\varepsilon}_t = \sum_{t=1}^{n} y_t - \hat{y}_t$$

As our objective is to generate a model that makes the smallest mistakes possible, a tempting way to proceed would be to minimise the previous equation. However, note that the residuals can be either positive or negative. Hence, if we would were to minimise the previous equation we would be making one inappropriate assumption. Namely, that models that make no mistake at all are as equally good as models with a huge amount of mistakes that counterbalance each other.

To see this, assume a candidate model A, with two observations ($t = 1{,}2$). This model has $\hat{\varepsilon}_{t=1} = 1{,}000$ and $\hat{\varepsilon}_{t=2} = -1{,}000$. Hence, it follows that $\sum_{t=1}^{n} \hat{\varepsilon}_t = (1{,}000 - 1{,}000) = 0$. Now consider another candidate model, B, with two observations ($t = 1{,}2$). This model has $\hat{\varepsilon}_{t=1} = 0$ and $\hat{\varepsilon}_{t=2} = 0$. Hence, it follows that $\sum_{t=1}^{n} \hat{\varepsilon}_t = (0 + 0) = 0$. If we were to minimise the sum of the residuals, it follows that we wouldn't be able to differentiate between models A and B. This is clearly a wrong approach, as we can see that model B is better than model A due to the fact it perfectly predicts our data.

One approach that we can take to solve this problem is to square the residuals and sum the square of the residuals. The rationale for this approach is that it solves the previous issue. To see this, note that the square of either a positive or a negative number is always a positive number. Hence, we get rid of the *counterbalancing issue* mentioned before. In the vein of our previous example, model A will have $(\hat{\varepsilon}_{t=1})^2 = 1{,}000^2 = 1{,}000{,}000$ and $(\hat{\varepsilon}_{t=2})^2 = (-1{,}000)^2 = 1{,}000{,}000$. Hence, $\sum_{t=1}^{n}(\hat{\varepsilon}_t)^2 = (1{,}000{,}000 + 1{,}000{,}000) = 2{,}000{,}000$ would be the sum of the residuals of model A. On the other hand, model B's squared residuals would be $(\hat{\varepsilon}_{t=1})^2 = 0^2 = 0$ and $(\hat{\varepsilon}_{t=2})^2 = (0)^2 = 0$. Hence, its sum of squared residuals would be $\sum_{t=1}^{n}(\hat{\varepsilon}_t)^2 = (0 + 0) = 0$. We can see that, when minimising the square of residuals, model B would appear as the clear winner. We can write the sum of squared residuals as:

$$S(\hat{\alpha}, \hat{\beta}) = \sum_{t=1}^{n}(\hat{\varepsilon}_t)^2 = \sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

Given that our prediction model is given by the equation $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$, we can substitute it in the previous equation to get:

$$S(\hat{\alpha}, \hat{\beta}) = \sum_{t=1}^{n}(\hat{\varepsilon}_t)^2 = \sum_{t=1}^{n}\left(y_t - \hat{\alpha} - \hat{\beta}x_t\right)^2$$

In order to find the Ordinary Least Squares estimators of α and β, we just need to solve an unconstrained minimisation problem (as we are not given any restrictions to complement the previous equation):

$$\min_{\hat{\alpha}, \hat{\beta}}\left[S(\hat{\alpha}, \hat{\beta}) = \sum_{t=1}^{n}\left(y_t - \hat{\alpha} - \hat{\beta}x_t\right)^2\right]$$

The solution to the above problem requires us to find the First Order Conditions (FOC's henceforth), which require setting the relevant first order derivatives equal to 0:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow 2\sum_{t=1}^{n} -\left(y_t - \hat{\alpha} - \hat{\beta}x_t\right) = 0$$

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow 2\sum_{t=1}^{n} -x_t * \left(y_t - \hat{\alpha} - \hat{\beta}x_t\right) = 0$$

Which collapse to:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow \sum_{t=1}^{n} -\left(y_t - \hat{\alpha} - \hat{\beta}x_t\right) = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow \sum_{t=1}^{n} -x_t * \left(y_t - \hat{\alpha} - \hat{\beta} x_t\right) = 0$$

We can expand both equations to get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow \sum_{t=1}^{n} -y_t + \hat{\alpha} + \hat{\beta} x_t = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow \sum_{t=1}^{n} -x_t y_t + \hat{\alpha} x_t + \hat{\beta} x_t x_t = 0$$

Using the relevant rule of working with the summation operator[6], we can rewrite the FOC's as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow -\sum_{t=1}^{n} y_t + \sum_{t=1}^{n} \hat{\alpha} + \sum_{t=1}^{n} \hat{\beta} x_t = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \hat{\alpha} x_t + \sum_{t=1}^{n} \hat{\beta} x_t x_t = 0$$

By taking the constant $\hat{\beta}$ outside of the summation operators[7], and by recalling the rule stating that the summation of a constant is n times the constant ($\sum_{t=1}^{n} c = nc$), we can rewrite the FOC's as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow -\sum_{t=1}^{n} y_t + n\hat{\alpha} + \hat{\beta} \sum_{t=1}^{n} x_t = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \hat{\alpha} x_t + \hat{\beta} \sum_{t=1}^{n} x_t x_t = 0$$

Isolating $n\hat{\alpha}$ in the Left Hand Side (LHS henceforth) of the first FOC, we get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow n\hat{\alpha} = \sum_{t=1}^{n} y_t - \hat{\beta} \sum_{t=1}^{n} x_t$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \hat{\alpha} x_t + \hat{\beta} \sum_{t=1}^{n} x_t x_t = 0$$

[6] $\sum_{i=0}^{n}(ax + bx) = \sum_{i=0}^{n} ax + \sum_{i=0}^{n} bx$
[7] Another rule regarding the summation operator is the following one: $\sum_{i=0}^{n} ax = a\sum_{i=0}^{n} x$

Isolating $\hat{\alpha}$ in the LFH of the first FOC, we get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\alpha}} = 0 \Rightarrow \hat{\alpha} = \frac{\sum_{t=1}^{n} y_t}{n} - \hat{\beta}\frac{\sum_{t=1}^{n} x_t}{n}$$

As, by definition, $\bar{z} = \frac{\sum_{t=1}^{n} z_t}{n}$, we can rewrite the previous expression as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\alpha}} = 0 \Rightarrow \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

Giving us the OLS estimator $\hat{\alpha}$ for the population parameter $\alpha.$ As we can see, the constant of our predicted model crucially depends on the averages of the independent and the dependent variables.

Now, substituting the expression for $\hat{\alpha}$ in the second FOC, we get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} (\bar{y} - \hat{\beta}\bar{x})x_t + \hat{\beta}\sum_{t=1}^{n} x_t x_t = 0$$

Expanding the parenthesis, we can rewrite the second FOC as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \bar{y}x_t - \hat{\beta}\bar{x}x_t + \hat{\beta}\sum_{t=1}^{n} x_t x_t = 0$$

As before, by noticing that $\sum_{i=0}^{n}(ax + bx) = \sum_{i=0}^{n} ax + \sum_{i=0}^{n} bx$, we can rewrite the previous FOC as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \bar{y}x_t - \sum_{t=1}^{n} \hat{\beta}\bar{x}x_t + \hat{\beta}\sum_{t=1}^{n} x_t x_t = 0$$

Taking the constant $\hat{\beta}$ outside of the summation term, we get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \bar{y}x_t - \hat{\beta}\sum_{t=1}^{n} \bar{x}x_t + \hat{\beta}\sum_{t=1}^{n} x_t x_t = 0$$

Taking $\hat{\beta}$ as a common factor, we can rewrite the previous expression as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t y_t + \sum_{t=1}^{n} \bar{y}x_t - \hat{\beta}\left(\sum_{t=1}^{n} \bar{x}x_t - \sum_{t=1}^{n} x_t x_t\right) = 0$$

Using x_t for both the terms inside and outside the parenthesis, and noticing that $\sum_{i=0}^{n} ax + \sum_{i=0}^{n} bx = \sum_{i=0}^{n}(ax + bx)$, we can rewrite the previous expression as:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow -\sum_{t=1}^{n} x_t(y_t - \bar{y}) + \hat{\beta}\left(\sum_{t=1}^{n} x_t(x_t - \bar{x})\right) = 0$$

Rearranging, we get:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow \hat{\beta}\left(\sum_{t=1}^{n} x_t(x_t - \bar{x})\right) = \sum_{t=1}^{n} x_t(y_t - \bar{y})$$

Isolating $\hat{\beta}$, we get:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow \hat{\beta} = \frac{\sum_{t=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} x_t(x_t - \bar{x})}$$

Hence, our estimators for the population parameters are:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = \frac{\sum_{t=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} x_t(x_t - \bar{x})}$$

(b) Show that $\widehat{\boldsymbol{\beta}}$ can be written as

$$\widehat{\boldsymbol{\beta}} = \sum_{t=1}^{n} \boldsymbol{c_t y_t}$$

And determine the coefficients $\boldsymbol{c_t}$ for $\boldsymbol{t = 1, 2, \dots, n}$.

In order to write $\hat{\beta}$ as the summation suggested above, we need to transform the expression so that we have y_t multiplied by another term. To do that, we subdivide this question into two parts. First, we transform the numerator and, second, we transform the denominator.

a. Transforming the numerator

The numerator of $\hat{\beta}$ is given by the following expression:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y})$$

We expand the previous expression to get:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - x_t \bar{y}$$

Using the rules of summation, we rewrite the previous equation as:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - \sum_{t=1}^{n} x_t \bar{y}$$

Taking the constant $\bar{y}$ outside of the second summation term, we get:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - \bar{y} \sum_{t=1}^{n} x_t$$

Substituting $\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}$ in the previous expression, we get:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - \frac{\sum_{t=1}^{n} y_t}{n} \sum_{t=1}^{n} x_t = \sum_{t=1}^{n} x_t y_t - \sum_{t=1}^{n} y_t \frac{\sum_{t=1}^{n} x_t}{n} = \sum_{t=1}^{n} x_t y_t - \sum_{t=1}^{n} y_t \, \bar{x}_t$$

Using the rules of summation outlined above, we can rewrite the two summation terms as a single summation term:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - \sum_{t=1}^{n} y_t \bar{x}_t = \sum_{t=1}^{n} x_t y_t - y_t \bar{x}_t$$

Using y_t as a common factor, we can rewrite the previous expression as:

$$\sum_{t=1}^{n} x_t(y_t - \bar{y}) = \sum_{t=1}^{n} x_t y_t - y_t \bar{x}_t = \sum_{t=1}^{n} y_t(x_t - \bar{x}_t)$$

Which is the transformed version of the numerator. As we can see, now we have the expression in terms of y_t and an extra term multiplying it.

b. Transforming the denominator

The numerator of $\hat{\beta}$ is given by the following expression:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x})$$

We expand the parenthesis within the previous expression to get:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t x_t - x_t \bar{x}$$

Applying the rules of the summation operator, we can divide the previous expression into two different summation terms to get:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t x_t - \sum_{t=1}^{n} x_t \bar{x} = \sum_{t=1}^{n} x_t^2 - \sum_{t=1}^{n} x_t \bar{x}$$

Noticing that $-\sum_{t=1}^{n} x_t \bar{x} = \sum_{t=1}^{n} x_t \bar{x} - 2\sum_{t=1}^{n} x_t \bar{x}$, we can rewrite the previous expression as:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t^2 + \sum_{t=1}^{n} x_t \bar{x} - 2\sum_{t=1}^{n} x_t \bar{x}$$

Taking the constant $\bar{x}$ out of the second summation term, we get:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t^2 + \bar{x}\sum_{t=1}^{n} x_t - 2\sum_{t=1}^{n} x_t\bar{x}$$

Noticing that $\bar{x} = \frac{\sum_{t=1}^{n} x_t}{n}$ and, hence, $n\bar{x} = \sum_{t=1}^{n} x_t$, we can substitute $\sum_{t=1}^{n} x_t$ in the second term in order to get:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t^2 + \bar{x}n\bar{x} - 2\sum_{t=1}^{n} x_t\bar{x}$$

Hence,

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t^2 + n\bar{x}^2 - 2\sum_{t=1}^{n} x_t\bar{x}$$

Noticing that $\bar{x}^2$ is a constant and applying the rule $\sum_{t=1}^{n} c = nc$, we get:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} x_t^2 + \sum_{t=1}^{n} \bar{x}^2 - 2\sum_{t=1}^{n} x_t\bar{x}$$

Using the identity $(a-b)^2 \equiv a^2 + b^2 - 2ab$, it is easy to see that the previous equation is equivalent to:

$$\sum_{t=1}^{n} x_t(x_t - \bar{x}) = \sum_{t=1}^{n} (x_t - \bar{x})^2$$

c. Rewriting $\hat{\beta}$ as $\hat{\beta} = \sum_{t=1}^{n} c_t y_t$

Using the results of parts *a.* and *b.*, we can rewrite the estimator $\hat{\beta}$ as:

$$\hat{\beta} = \frac{\sum_{t=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} x_t(x_t - \bar{x})} = \frac{\sum_{t=1}^{n} y_t(x_t - \bar{x}_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

Hence, it follows that $c_t = \frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$ describes the coefficients asked for. To see this, just substitute the formula for c_t in the estimator of $\hat{\beta}$ to get:

$$\hat{\beta} = \sum_{t=1}^{n} c_t y_t = \sum_{t=1}^{n} \frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} y_t$$

Once we apply the summation in the denominator, we end up with a constant, which can be taken out of the global summation term to get:

$$\hat{\beta} = \sum_{t=1}^{n} c_t y_t = \frac{\sum_{t=1}^{n}(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} y_t = \frac{\sum_{t=1}^{n}(x_t - \bar{x}) y_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

(c) Using (1) show that the variance of $\widehat{\boldsymbol{\beta}}$ is given by

$$\boldsymbol{V}[\widehat{\boldsymbol{\beta}}] = \frac{\boldsymbol{\sigma}^2}{\sum_{i=1}^{n}(\boldsymbol{x}_t - \overline{\boldsymbol{x}})^2}$$

And explain why empirical investigation might be hampered if there were only small variations in the values of the independent variable. What would be the consequence if $\boldsymbol{x}_t = \boldsymbol{x}_s$ for every $\boldsymbol{t}$ and $\boldsymbol{s}$?

Given that $\hat{\beta} = \sum_{t=1}^{n} c_t y_t$, it follows that:

$$V[\hat{\beta}] = V\left[\sum_{t=1}^{n} c_t y_t\right]$$

Using the rules of variance, and given that c_t is a constant and that the y_t's are independent, we can rewrite the expression as:

$$V[\hat{\beta}] = \sum_{t=1}^{n} V[c_t y_t] = \sum_{t=1}^{n} c_t^2 V[y_t]$$

Substituting the expression for y_t ($y_t = \alpha + \beta x_t + \varepsilon_t$) in the variance, we get the following expression:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 V[\alpha + \beta x_t + \varepsilon_t]$$

Knowing that the generic form of the variance is $V[y_t] = E[(y_t - E[y_t])^2]$, we can rewrite the previous expression as:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 E[(\alpha + \beta x_t + \varepsilon_t - E[\alpha + \beta x_t + \varepsilon_t])^2]$$

Using the rules of expectations, we can divide the $E[\alpha + \beta x_t + \varepsilon_t]$ term into three terms to get:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 E[(\alpha + \beta x_t + \varepsilon_t - E[\alpha] - E[\beta x_t] - E[\varepsilon_t])^2]$$

By noticing that $E[\alpha] = \alpha$, $E[\beta x_t] = \beta x_t$[8] and that $E[\varepsilon_t] = 0$, we can rewrite the previous expression as:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 E[(\alpha + \beta x_t + \varepsilon_t - \alpha - \beta x_t - 0)^2]$$

Which collapses to:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 E[(\varepsilon_t)^2]$$

Given that we are told that $\varepsilon_t \sim IIDN(0, \sigma^2)$, then it follows that:

$$V[\varepsilon_t] = E[(\varepsilon_t - E[\varepsilon_t])^2] = E[(\varepsilon_t - 0)^2] = E[\varepsilon_t^2] = \sigma^2$$

Hence, substituting $E[\varepsilon_t^2] = \sigma^2$ in the expression for the variance of $\hat{\beta}$, we get:

$$V[\hat{\beta}] = \sum_{t=1}^{n} c_t^2 \sigma^2$$

As σ^2 is a constant, we can take it out of the summation term to get:

$$V[\hat{\beta}] = \sigma^2 \sum_{t=1}^{n} c_t^2$$

Given that $c_t = \frac{(x_t - \bar{x}_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$, the previous expression becomes:

$$V[\hat{\beta}] = \sigma^2 \sum_{t=1}^{n} \left(\frac{(x_t - \bar{x}_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \right)^2$$

Which, after applying some algebraic calculations, can be rewritten as:

$$V[\hat{\beta}] = \sigma^2 \frac{\sum_{t=1}^{n}(x_t - \bar{x}_t)^2}{(\sum_{t=1}^{n}(x_t - \bar{x})^2)^2} = \sigma^2 \frac{\left(\sum_{t=1}^{n}(x_t - \bar{x}_t)^2\right)^1}{(\sum_{t=1}^{n}(x_t - \bar{x})^2)^2} = \frac{\sigma^2}{\sum_{t=1}^{n}(x_t - \bar{x}_t)^2}$$

---

[8] Notice that our population model $y_t = \alpha + \beta x_t + \varepsilon_t$ only assumes ε_t to be a random variable. Both α, β and x_t are not random but known variables and parameters and, hence, their expected value will just be the expected value of a constant. It, then, follows that the variance of y_t, as we shall show later on, will be directly proportional to the variance of ε_t.

If there were only small variations in the values of the independent variable, then our investigation might be hampered. To see this, we need to show that as the variation of the values of the independent variable decrease, the variance of our estimator $\hat{\beta}$ increases. This means more noise in the data and, as a consequence, a greater difficulty in achieving statistical significance for a given value of $\hat{\beta}$. To see this, let's define an arbitrarily small value as δ. To see how the variance reacts to small variations in the values of the independent variable, we compute the limit of the variance when $x_t - \bar{x}_t$ converges to the defined arbitrarily small value:

$$\lim_{x_t - \bar{x}_t \to \delta} \left(\frac{\sigma^2}{\sum_{t=1}^{n} (x_t - \bar{x}_t)^2} \right) = \frac{\sigma^2}{\sum_{t=1}^{n} \delta^2} = \frac{\sigma^2}{n\delta^2}$$

Now, by taking the derivative of the previous expression with respect to δ we will see how the variance varies when the variation of the values of the independent values decrease:

$$\frac{d\left(\frac{\sigma^2}{n\delta^2}\right)}{d\delta} = -\frac{2n\sigma^2\delta}{n^2\delta^4} = -\frac{2\sigma^2}{n\delta^3} < 0$$

As we can see, when δ gets bigger, the variance gets smaller. It, then, follows that when δ gets smaller the variance gets bigger. Hence, when the variation of the values of the independent variable are small, the variance will tend to be big. Also, it is easy to see that, when $x_t = x_s$, then

$$\bar{x} = \frac{\sum_{t=1}^{n} x_t}{n} = \frac{x_t + x_t + \cdots + x_t}{n} = \frac{nx_t}{n} = x_t$$

And, substituting $\bar{x} = x_t$ in the formula for the variance of $\hat{\beta}$ found before, we get:

$$V[\hat{\beta}]\big|_{x_t = \bar{x}} = \frac{\sigma^2}{\sum_{t=1}^{n} (x_t - x_t)^2} = \frac{\sigma^2}{\sum_{t=1}^{n} (0)^2} = \frac{\sigma^2}{0} = \infty$$

Hence, the consequence of $x_t = x_s$ for every t and s would be an infinitely large variance. Thus, no statistical significance of the independent variable could ever be reached.

(d) Let $\boldsymbol{d_1, \dots, d_n}$ be non-random coefficients that satisfy the following;

$$\boldsymbol{\sum_{i=1}^{n} d_t = 0} \quad \textbf{and} \quad \boldsymbol{\sum_{i=1}^{n} d_t x_t = 1}$$

Show that the estimator defined by

$$\boldsymbol{\tilde{\beta}_d = \sum_{i=1}^{n} d_t y_t}$$

Is an unbiased estimator of $\boldsymbol{\beta}$.

Taking expectations, we get:

$$E[\tilde{\beta}_d] = E\left[\sum_{i=1}^{n} d_t y_t\right]$$

As the y_t's are independent, we can rewrite the previous equation as:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} E[d_t y_t]$$

Substituting $y_t = \alpha + \beta x_t + \varepsilon_t$ into the previous expression, we get:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} E[d_t(\alpha + \beta x_t + \varepsilon_t)]$$

Expanding the parenthesis, the previous expression becomes:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} (E[d_t\alpha] + E[d_t\beta x_t] + E[d_t\varepsilon_t])$$

Using the rules of the summation operator, we can divide the Right Hand Side into three summations:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} E[d_t\alpha] + \sum_{i=1}^{n} E[d_t\beta x_t] + \sum_{i=1}^{n} E[d_t\varepsilon_t]$$

Which, given that $E[d_t] = d_t$, becomes:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} d_t \alpha + \sum_{i=1}^{n} d_t \beta x_t + \sum_{i=1}^{n} d_t E[\varepsilon_t]$$

As $E[\varepsilon_t] = 0$, the third term vanishes:

$$E[\tilde{\beta}_d] = \sum_{i=1}^{n} d_t \alpha + \sum_{i=1}^{n} d_t \beta x_t$$

Also, as α and β are constants, we can take them out of the summation to get:

$$E[\tilde{\beta}_d] = \alpha \sum_{i=1}^{n} d_t + \beta \sum_{i=1}^{n} d_t x_t$$

And, given that $\sum_{i=1}^{n} d_t = 0$ and $\sum_{i=1}^{n} d_t x_t = 1$ as assumed in the question, the previous equation becomes:

$$E[\tilde{\beta}_d] = \alpha(0) + \beta(1) = \beta$$

Hence, if the assumptions of the question hold true, then the estimator $\tilde{\beta}_d$ is unbiased as

$$E[\tilde{\beta}_d] = \beta$$

(e) Show that the variance of the unbiased estimator $\tilde{\beta}_d$ is minimised when

$$d_t = c_t \quad \text{for all } t$$

Where the coefficients c_t are as determined in part (c).

Showing that the variance of the estimator $\tilde{\beta}_d$ is minimised when $d_t = c_t$ implies showing that the variance of the OLS estimator $\hat{\beta}$ is the minimum possible one. Hence, the proved we are asked for is the proof of the Gauss-Markov theorem. To start, and similarly to part (d),

$$V[\tilde{\beta}_d] = V\left[\sum_{t=1}^{n}[d_t y_t]\right] = \sum_{t=1}^{n}[V[d_t y_t]] = \sum_{t=1}^{n}[d_t^2 V[y_t]] = \sigma^2 \sum_{t=1}^{n}[d_t^2]$$

Now, we want to manipulate the expression so that it depends on both c_t and d_t. Later on, we then proceed to show that the term involving d_t is positive and, hence, the variance is minimised when $d_t = c_t$.

Given that $c_t - c_t = 0$, we can rewrite the previous equation as:

$$V[\tilde{\beta}_d] = \sigma^2 \sum_{t=1}^{n}[(c_t - c_t + d_t)^2]$$

Which can be rearranged as:

$$V[\tilde{\beta}_d] = \sigma^2 \sum_{t=1}^{n}\left[\left(c_t + (d_t - c_t)\right)^2\right]$$

Expanding, we get:

$$V[\tilde{\beta}_d] = \sigma^2 \sum_{t=1}^{n}[c_t^2 + (d_t - c_t)^2 + 2\left(c_t(d_t - c_t)\right)]$$

Using the rules of the summation operator, we can divide the RHS into three summation terms:

$$V[\tilde{\beta}_d] = \sigma^2\left[\sum_{t=1}^{n} c_t^2 + \sum_{t=1}^{n}(d_t - c_t)^2 + \sum_{t=1}^{n} 2c_t(d_t - c_t)\right]$$

The next step is to show that the third element within the brackets is equal to 0. To see this, let's just write the third element in the brackets below:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\sum_{t=1}^{n} c_t(d_t - c_t)$$

Substituting the first c_t by the formula we previously found for c_t ($c_t = \frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$), we get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\sum_{t=1}^{n}\left(\frac{(x_t - \bar{x}_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)(d_t - c_t) = 2\left(\frac{\sum_{t=1}^{n}(x_t - \bar{x})(d_t - c_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Expanding the numerator, we get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{\sum_{t=1}^{n}(x_t d_t - x_t c_t - \bar{x} d_t + \bar{x} c_t)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Using the rules of the summation operator, we can divide the numerator into four summation terms:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{\sum_{t=1}^{n} x_t d_t - \sum_{t=1}^{n} x_t c_t - \sum_{t=1}^{n} \bar{x} d_t + \sum_{t=1}^{n} \bar{x} c_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

We can take the constant $\bar{x}$ out of the third summation term to get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{\sum_{t=1}^{n} x_t d_t - \sum_{t=1}^{n} x_t c_t - \bar{x}\sum_{t=1}^{n} d_t + \sum_{t=1}^{n} \bar{x} c_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

As $\sum_{t=1}^{n} x_t d_t = 1$ and $\sum_{t=1}^{n} d_t = 0$ because of the assumptions made at question (d), the previous equation becomes:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n} x_t c_t - \bar{x}(0) + \sum_{t=1}^{n} \bar{x} c_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Which simplifies to:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n} x_t c_t + \sum_{t=1}^{n} \bar{x} c_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Now, substituting the remaining c_t's by $c_t = \frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$, we get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n} x_t \left(\frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) + \sum_{t=1}^{n} \bar{x}\left(\frac{(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Expanding the second and third terms of the global parenthesis, we get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n}\left(\frac{x_t^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2} - \frac{x_t\bar{x}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) + \sum_{t=1}^{n}\left(\frac{\bar{x}x_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2} - \frac{\bar{x}^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Using the rules of the summation operator, we can divide each of the two summation terms of the numerator into two summation terms to get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n}\left(\frac{x_t^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) + \sum_{t=1}^{n}\left(\frac{x_t\bar{x}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) + \sum_{t=1}^{n}\left(\frac{\bar{x}x_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) - \sum_{t=1}^{n}\left(\frac{\bar{x}^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

We can, now, group the four summation terms of the numerator together to get:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 + \sum_{t=1}^{n}\left(-\frac{x_t^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2} + \frac{x_t\bar{x}}{\sum_{t=1}^{n}(x_t - \bar{x})^2} + \frac{\bar{x}x_t}{\sum_{t=1}^{n}(x_t - \bar{x})^2} - \frac{\bar{x}^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Which can be rewritten as:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \sum_{t=1}^{n}\left(\frac{x_t^2 + \bar{x}^2 - 2x_t\bar{x}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) = 2\left(\frac{1 - \frac{\sum_{t=1}^{n} x_t^2 + \bar{x}^2 - 2x_t\bar{x}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

As $(a - b)^2 \equiv a^2 + b^2 - 2ab$, we can rewrite the previous expression as:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - \frac{\sum_{t=1}^{n}(x_t - \bar{x})^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right)$$

Which, when simplifying, becomes:

$$\sum_{t=1}^{n} 2c_t(d_t - c_t) = 2\left(\frac{1 - 1}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) = 2\left(\frac{0}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right) = 2(0) = 0$$

Hence, as we have demonstrated that the third element within the brackets of the expression for the variance of the estimator $\tilde{\beta}_d$ is equal to 0, we can now rewrite the expression for the variance of the estimator $\tilde{\beta}_d$ as:

$$V[\tilde{\beta}_d] = \sigma^2 \left[\sum_{t=1}^{n} c_t^2 + \sum_{t=1}^{n} (d_t - c_t)^2 \right]$$

Now, as a given number squared is always positive regardless of its sign, then it follows that $(d_t - c_t)^2 \geq 0$ and, consequently,

$$\sum_{t=1}^{n} (d_t - c_t)^2 \geq 0$$

Hence, the variance of the estimator $\tilde{\beta}_d$ can be written more generically as:

$$V[\tilde{\beta}_d] = \begin{cases} \sigma^2 \sum_{t=1}^{n} c_t^2 \;\; if\ d_t = c_t \\ \sigma^2 \left[\sum_{t=1}^{n} c_t^2 + \geq 0 \right] \; if\ d_t \neq c_t \end{cases}$$

Thus, the variance of the estimator $\tilde{\beta}_d$ is minimised whenever $c_t = d_t$. Or, in other terms, **the variance of the estimator $\boldsymbol{\tilde{\beta}_d}$ is minimised whenever $\boldsymbol{\tilde{\beta}_d}$ is the OLS estimator of $\boldsymbol{\beta}$.**

(f) Explain why both $\hat{\alpha}$ and $\hat{\beta}$ are consistent estimators.

We know that an estimator is consistent provided that

$$\lim_{n\to\infty} \Pr\left[\left|\hat{\beta}-\beta\right| \geq \varepsilon\right] = 0$$

In addition, Chebyshev's inequality states the following:

$$\Pr\left[\left|\hat{\beta}-E\left[\hat{\beta}\right]\right| \geq \varepsilon\right] \leq \frac{V\left[\hat{\beta}\right]}{\varepsilon^2}$$

Provided that $\hat{\beta}$ is an unbiased estimator, then the first expression is equivalent to

$$\lim_{n\to\infty} \Pr\left[\left|\hat{\beta}-E\left[\hat{\beta}\right]\right| \geq \varepsilon\right] = 0$$

Equalling the lower bound of Chebyshev's inequality with the equality of the previous expression implies:

$$\frac{V\left[\hat{\beta}\right]}{\varepsilon^2} = 0 \Rightarrow V\left[\hat{\beta}\right] = 0$$

As n tends to ∞. Given that the expression for the variance of $\hat{\beta}$ was found before:

$$V\left[\hat{\beta}\right] = \frac{\sigma^2}{\sum_{t=1}^{n}\left(x_t-\bar{x}\right)^2}$$

It follows that:

$$\lim_{n\to\infty} V\left[\hat{\beta}\right] = \lim_{n\to\infty} \frac{\sigma^2}{\sum_{t=1}^{n}\left(x_t-\bar{x}\right)^2} = 0$$

As n tends to ∞, the denominator will be composed of infinite terms:

$$\lim_{n\to\infty} V\left[\hat{\beta}\right] = \frac{\sigma^2}{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots}$$

If all x_t's are equal, then it follows that $x_1 = x_2 = \cdots = \bar{x}$ and, as a result, the denominator will be equal to 0. However, if, for some t and $x_t \neq x_s$, then the denominator will get big. More formally, let's define $\rho \in [0,1]$ as the percentage of observations x_t different from $\bar{x}$. Then, it follows that, as n tends to ∞, the denominator will have $\rho * \infty = \infty$ positive terms (as any number squared is positive. Hence, we can rewrite the previous expression as:

$$\lim_{n\to\infty} V\left[\hat{\beta}\right] = \frac{\sigma^2}{(>0) + (>0) + \cdots} = \frac{\sigma^2}{\infty} = 0$$

A similar argument can be applied to the variance of the estimator $\hat{\alpha}$.

L11MEE TUTORIALS – SEMESTER 2 – TUTORIAL 5

By ERNESTO M. GAVASSA PEREZ*

*This document was compiled based on the answers provided by Professor Patrick Marsh. Hence, all the credit should be given to him

1. A classical regression model is written,

$$\boldsymbol{y_t = \alpha + \beta x_t + \varepsilon_t, \quad \varepsilon_t \sim IIDN(0, \sigma^2), \quad t = 1, 2, \dots, n}$$

Where the values of $\boldsymbol{x_1, \dots, x_n}$ are not random. Let $\boldsymbol{\hat{\alpha}}$ and $\boldsymbol{\hat{\beta}}$ denote the OLS estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, with the fitted values $\boldsymbol{\hat{y}_t}$ and residuals $\boldsymbol{\hat{\varepsilon}_t}$ given by;

$$\boldsymbol{\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t \quad \& \quad \hat{\varepsilon}_t = y_t - \hat{y}_t.}$$

(a) Define $\boldsymbol{TSS = \sum_{i=1}^{n}(y_t - \bar{y})^2}$, $\boldsymbol{ESS = \sum_{i=1}^{n}(\hat{y}_t - \bar{y})^2}$ and $\boldsymbol{RSS = \sum_{i=1}^{n} \hat{\varepsilon}_t^2}$,

Then show that

$$\boldsymbol{TSS = ESS + RSS}$$

As stated in the question, the total sum of squares is defined to be the square of the deviation of the dependent variable around its mean:

$$TSS = \sum_{i=1}^{n}(y_t - \bar{y})^2$$

Given that the residuals are assumed to be $\hat{\varepsilon}_t = y_t - \hat{y}_t$, we can isolate the dependent variable to get the following identity:

$$\hat{y}_t + \hat{\varepsilon}_t = y_t$$

Substituting y_t in the TSS formula, we get:

$$TSS = \sum_{i=1}^{n}(y_t - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_t + \hat{\varepsilon}_t - \bar{y})^2$$

Which, after rearranging, looks like:

$$TSS = \sum_{i=1}^{n}\left((\hat{y}_t - \bar{y}) + \hat{\varepsilon}_t\right)^2$$

Expanding, we get:

$$TSS = \sum_{i=1}^{n}(\hat{y}_t - \bar{y})^2 + \hat{\varepsilon}_t^2 - 2\hat{\varepsilon}_t(\hat{y}_t - \bar{y})$$

Using the rules of the summation operator, we can divide the summation into three different summations to get:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - \sum_{i=1}^{n} 2\hat{\varepsilon}_t(\hat{y}_t - \bar{y})$$

Using the rules of the summation operator, we can take the 2 outside the third summation operator; and, additionally expanding the parenthesis of the third element, we get:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\sum_{i=1}^{n} \hat{\varepsilon}_t\hat{y}_t - \hat{\varepsilon}_t\bar{y}$$

We can divide the third term into two summations:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\sum_{i=1}^{n} \hat{\varepsilon}_t\hat{y}_t + 2\sum_{i=1}^{n} \hat{\varepsilon}_t\bar{y}$$

By noticing that $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$, we can substitute the third element to get:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\sum_{i=1}^{n} \hat{\varepsilon}_t(\hat{\alpha} + \hat{\beta}x_t) + 2\sum_{i=1}^{n} \hat{\varepsilon}_t\bar{y}$$

Again, expanding the third element and dividing it into two summation terms, we get:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\sum_{i=1}^{n} \hat{\varepsilon}_t\hat{\alpha} - 2\sum_{i=1}^{n} \hat{\varepsilon}_t\hat{\beta}x_t + 2\sum_{i=1}^{n} \hat{\varepsilon}_t\bar{y}$$

As $\hat{\alpha}$, $\hat{\beta}$ and $\bar{y}$ are constants, they can be taken out of the summation terms:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\hat{\alpha}\sum_{i=1}^{n} \hat{\varepsilon}_t - 2\hat{\beta}\sum_{i=1}^{n} \hat{\varepsilon}_t x_t + 2\bar{y}\sum_{i=1}^{n} \hat{\varepsilon}_t$$

One should notice, at this point, that $\sum_{i=1}^{n} \hat{\varepsilon}_t = 0$ and $\sum_{i=1}^{n} \hat{\varepsilon}_t x_t = 0$. To see this, just rewrite the FOC's of the OLS we covered in the previous tutorial:

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\alpha}} = 0 \Rightarrow 2\sum_{t=1}^{n} -(y_t - \hat{\alpha} - \hat{\beta}x_t) = 0$$

$$\frac{\partial S(\hat{\alpha}, \hat{\beta})}{\partial \hat{\beta}} = 0 \Rightarrow 2\sum_{t=1}^{n} -x_t * (y_t - \hat{\alpha} - \hat{\beta}x_t) = 0$$

Now, given that $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$ as commented earlier, the previous FOC's can be rewritten as:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\alpha}} = 0 \Rightarrow 2\sum_{t=1}^{n} -\left(y_t - (\hat{\alpha} + \hat{\beta}x_t)\right) = 2\sum_{t=1}^{n} -(y_t - \hat{y}_t) = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow 2\sum_{t=1}^{n} -x_t * \left(y_t - (\hat{\alpha} + \hat{\beta}x_t)\right) = 2\sum_{t=1}^{n} -x_t * (y_t - \hat{y}_t) = 0$$

Recall that the question stated that $\hat{\varepsilon}_t = y_t - \hat{y}_t$. Substituting $\hat{\varepsilon}_t$ in the FOC's, we get:

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\alpha}} = 0 \Rightarrow 2\sum_{t=1}^{n} -(\hat{\varepsilon}_t) = -2\sum_{t=1}^{n} \hat{\varepsilon}_t = 0 \Rightarrow \sum_{t=1}^{n} \hat{\varepsilon}_t = 0$$

$$\frac{\partial S(\hat{\alpha},\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow 2\sum_{t=1}^{n} -x_t * \hat{\varepsilon}_t = -2\sum_{t=1}^{n} x_t * \hat{\varepsilon}_t = 0 \Rightarrow \sum_{t=1}^{n} x_t * \hat{\varepsilon}_t$$

Hence, by substituting $\sum_{i=1}^{n} \hat{\varepsilon}_t = 0$ and $\sum_{i=1}^{n} \hat{\varepsilon}_t x_t = 0$ in the TSS equation, we get:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 - 2\hat{\alpha}(0) - 2\hat{\beta}(0) + 2\bar{y}(0)$$

Which collapses to:

$$TSS = \sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2 + \sum_{i=1}^{n} \hat{\varepsilon}_t^2 = ESS + RSS$$

(b) Detail the role played by the coefficient of determination R^2 is assessing the overall fit of the fitted model.

The coefficient of determination R^2 is defined to be the deviation of the independent variable around its mean that is explained by the variation of the predicted variable $\hat{y}_t$ around its mean. Put shortly, R^2 is the amount of variance of the independent variable that can be predicted by our estimation. Notice that, as $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$, then $E[\hat{y}_t] = \bar{\hat{y}} = \alpha + \beta x_t = E[y_t] = \bar{y}$ as far as the OLS estimators are unbiased (they indeed are. See part (d) of tutorial 4 and the lecture notes).

Hence, we can rewrite ESS as:

$$ESS = \sum_{i=1}^{n} \left(\hat{y}_t - \bar{\hat{y}}\right)^2$$

Given the definition stated above, it becomes clear that $\frac{ESS}{TSS}$ is the coefficient of determination:

$$R^2 = \frac{ESS}{TSS}$$

As $TSS = ESS + RSS$, then, we can substitute TSS above to get:

$$R^2 = \frac{ESS}{ESS + RSS}$$

And, as $RSS = \sum_{i=1}^{n} \hat{\varepsilon}_t^2 \geq 0$ and $ESS = \sum_{i=1}^{n}\left(\hat{y}_t - \bar{\hat{y}}\right)^2 \geq 0$, as both are sums of squared numbers, it follows that

$$R^2 = \frac{\geq 0}{\geq 0} = (\geq 0)$$

Additionally,

$$R^2 = \frac{ESS}{ESS + \geq 0} = (\leq 1)$$

Hence, R^2 can only take values in the range 0 to 1. In other words, the percentage of variance of the independent variable that can be explained by our model lies within the conventional 0% - 100% limit:

$$R^2 \in [0,1]$$

Alternatively, one can divide the identity $TSS = ESS + RSS$ and divide it by TSS to get:

$$\frac{TSS}{TSS} = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

As we know that $R^2 = \frac{ESS}{TSS}$, substituting it and isolating R^2, we get:

$$1 - \frac{RSS}{TSS} = R^2$$

(c) Detail why the statistic

$$\boldsymbol{T_\beta} = \left(\frac{\boldsymbol{\hat{\beta} - \beta}}{\sqrt{\frac{\boldsymbol{\hat{\sigma}^2}}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}} \right) \sim \boldsymbol{t_{(n-2)}},$$

Where $\boldsymbol{t_{(n-2)}}$ has a student-t distribution with $\boldsymbol{n-2}$ degrees of freedom.

As discussed in the previous tutorial, $\hat{\beta} \sim N\left[\beta, \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}\right]$. Standardising the OLS estimator will yield a variable following a normal distribution with mean 0 and variance 1. Even when proved in previous tutorials, let's just prove it again so that we are confident enough with the suggested result. Taking expectations of the standardised OLS variable, we get:

$$E\left[\frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right]$$

Using the rules of expectations, we can divide the previous expression into two terms:

$$E\left[\frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right] = E\left[\frac{\hat{\beta}}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right] - E\left[\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right]$$

Given that x_t is not random, and that σ^2 and β are constant, we can use the rules of expectations to get:

$$E\left[\frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right] = \frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}} E[\hat{\beta}] - \frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}$$

Given that the OLS estimator is unbiased, then $E[\hat{\beta}] = \beta$. Substituting in the previous equation, we get:

$$E\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\beta-\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}=\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}-\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}=0$$

We can, now, express the variance of the standardised variable as follows:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=E\left[\left(\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}-E\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]\right)^2\right]$$

Given that $E\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=0$, as found before, the previous expression collapses to:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=E\left[\left(\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\right]$$

Expanding, we get:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=E\left[\left(\frac{\hat{\beta}}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2+\left(\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2-2\left(\frac{\hat{\beta}}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)\left(\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)\right]$$

Applying the rules of calculus, we can divide the previous expression into 3 terms:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=E\left[\left(\frac{\hat{\beta}}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\right]+E\left[\left(\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\right]+E\left[-2\left(\frac{\hat{\beta}}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)\left(\frac{\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)\right]$$

Given that x_t is not random, and that σ^2 and β are constant, we can use the roles of expectations to get:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2 E[\hat{\beta}^2]+\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2 E[\beta^2]-2\beta\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2 E[\hat{\beta}]$$

Taking $\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2$ as a common factor, we get:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\left(E[\hat{\beta}^2]+E[\beta^2]-2\beta E[\hat{\beta}]\right)$$

As β^2 is constant, we can take it out of the expectation. Also, as $\hat{\beta}$ is unbiased, then $E[\hat{\beta}]=\beta$, making the previous expression equivalent to:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\left(E[\hat{\beta}^2]+\beta^2-2\beta\beta\right)$$

Which collapses to:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\left(\frac{1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2\left(E[\hat{\beta}^2]+\beta^2-2\beta^2\right)$$

And further simplifies to:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right]=\left(\frac{1}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}\right)\left(E[\hat{\beta}^2]-\beta^2\right)$$

We do not know $E[\hat{\beta}^2]$, but we do know that:

$$V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} = E\left[\left(\hat{\beta} - \beta\right)^2\right]$$

Expanding the RHS, we get:

$$V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} = E\left[\hat{\beta}^2 + \beta^2 - 2\beta\hat{\beta}\right]$$

Which, using the rules of expectations, is equivalent to:

$$V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} = E\left[\hat{\beta}^2\right] + \beta^2 - 2\beta E\left[\hat{\beta}\right]$$

Again, as $\hat{\beta}$ is unbiased, we can substitute $E\left[\hat{\beta}\right] = \beta$ in the previous equation to get:

$$V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} = E\left[\hat{\beta}^2\right] + \beta^2 - 2\beta\beta = E\left[\hat{\beta}^2\right] + \beta^2 - 2\beta^2$$

Which simplifies to

$$V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} = E\left[\hat{\beta}^2\right] - \beta^2$$

Recalling that the expression for the variance of our standardised variable was:

$$V\left[\frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right] = \left(\frac{1}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}\right)\left(E\left[\hat{\beta}^2\right] - \beta^2\right)$$

And noticing that we just found $E\left[\hat{\beta}^2\right] - \beta^2 = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}$, we can substitute the expression for $E\left[\hat{\beta}^2\right] - \beta^2$ in the previous equation to get:

$$V\left[\frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}\right] = \left(\frac{1}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}\right)\left(\frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}\right)$$

Which simplifies to:

$$V\left[\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right] = \frac{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}} = 1$$

Hence, we have shown that $\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}} \sim N(0,1)$

More importantly, notice that, now, the standardised variable will give us a number. Also, given the confidence level we want to achieve, the statistical table of the normal distribution will give us a number against to which we can compare our statistic.

Or is it? **Notice that we have σ^2 in the denominator, which is a population parameter and, hence, unknown.** $\hat{\beta}$ is clearly known, as it is our OLS estimator; and $\sum_{i=1}^{n}(x_t-\bar{x})^2$ is also known as we observe our independent variable x_t. If we want to test whether $\hat{\beta}$ is statistically significantly different from an ex-ante hypothesized value of β, then we *don't care* about the real value of β but about our hypothesised value; so we could say we *know* β (**italics are crucial to understand this bit.** Of course we care about β, but in the domain of hypothesis testing our main focus is on the difference between the value we observe, $\hat{\beta}$, and the value we ex-ante hypothesize it should have). The main punchline in here is to note that $\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}$ is not a valid test statistic, as it would need us to know the population parameter σ^2, which we do not know.

Here, the use of the distributions we have studied in the lecture notes become massively important. The question that follows is: can we transform our normal (0,1) into another variable, following a distribution we know, that can allow us to test for the difference between $\hat{\beta}$ and β?

Instead of directly giving the answer, I believe to be useful to explain why we do actually need a t-student distribution. To see this, notice that a chi square distribution wouldn't solve our problem. If we square our standardised variable, we'd get the following:

$$\left(\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}\right)^2 = \frac{\left(\hat{\beta}-\beta\right)^2}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}} \sim \chi^2_{(1)}$$

As we know from the lecture notes, a normal (0,1) squared follows a chi quare distribution. In this case, as there is only one normal (0,1) squared, the previous expression involves a variable following a chi square distribution with 1 degree of freedom. However, and more crucially, we'd still be facing the same problem: namely, that we still have an unknown parameter (σ^2) affecting our variable. Therefore, just a chi square won't be able to suit our needs. An F-snedecor distribution may solve our issues. But, as we'll see later on, under some circumstances the square of the T statistic follows an F distribution.

Let's turn our focus to the remaining distribution that can be constructed from a normal (0,1): the t-student distribution. If we recall the lectures, we know that a t-student distribution is generated in the following way:

$$T = \frac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}} \sim t_{(n)}$$

Even when we do not know σ^2, we do actually know $\hat{\sigma}^2$. Hence, **if there was a way to generate a chi square variable with both $\boldsymbol{\sigma^2}$ and $\boldsymbol{\hat{\sigma}^2}$, such that the $\boldsymbol{\sigma^2}$'s would cancel out, we could use a t-statistic to test for the difference between the OLS estimator of $\boldsymbol{\beta}$ and $\boldsymbol{\beta}$ itself. Indeed**, it will be shown in later courses that:

$$\boldsymbol{\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-2)}}$$

Hence, it is the case that:

$$T = \frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}} \sim t_{(n-2)}$$

We can simplify by noticing that the $(n-2)$'s cancel out:

$$T=\frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}}=\frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}}\sim t_{(n-2)}$$

Which is equivalent to[9]:

$$T=\frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}}=\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}*\frac{\sqrt{\sigma^2}}{\sqrt{\hat{\sigma}^2}}\sim t_{(n-2)}$$

We can further write it as:

$$T=\frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}}=\frac{\left(\hat{\beta}-\beta\right)*\sqrt{\sigma^2}}{\sqrt{\sigma^2}\sqrt{\frac{1}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}\sqrt{\hat{\sigma}^2}}\sim t_{(n-2)}$$

From which it becomes clear that the σ^2's cancel out:

$$T=\frac{\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}}{\sqrt{\frac{\frac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}}=\frac{\left(\hat{\beta}-\beta\right)}{\sqrt{\frac{1}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}\sqrt{\hat{\sigma}^2}}\sim t_{(n-2)}$$

Again, using the rules of calculus (now we exploit the fact that $\sqrt{ab}=\sqrt{a}*\sqrt{b}$), we get:

[9] Remember that $\frac{\frac{a}{b}}{\frac{c}{c}}=\frac{a}{b}*\frac{d}{c}$ and that $\sqrt{\frac{a}{b}}=\frac{\sqrt{a}}{\sqrt{b}}$

$$T = \frac{\dfrac{\hat{\beta} - \beta}{\sqrt{\dfrac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}}{\sqrt{\dfrac{\dfrac{(n-2)\hat{\sigma}^2}{\sigma^2}}{n-2}}} = \frac{\left(\hat{\beta} - \beta\right)}{\sqrt{\dfrac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}} \sim t_{(n-2)}$$

Now, we have a very similar variable from the standardised one we started with. However, notice that the unique difference is that we now have $\hat{\sigma}^2$, which we know, instead of the unknown σ^2. This changes the distribution from a normal (0,1) to a t-student distribution with $(n-2)$ degrees of freedom. We can, now, compare the value of our statistic with the critical value we get from the t-student distribution given the n of our sample and the confidence level we want to use. Furthermore, the additional fact that $\hat{\sigma}^2$ is an unbiased estimator of the population parameter σ^2 makes it even more attractive.

(d) A researcher has data from 27 households and is interested in the relationship between weekly household consumption Y and weekly household income X. The results of their OLS estimation are

$$\boldsymbol{y_t = 5.0 + 0.25x_t + \hat{\varepsilon}_t, \quad R^2 = 0.5, \quad n = 27}$$

$$\boldsymbol{(1.0) \quad (0.05).}$$

Where $\hat{\varepsilon}_t$ denotes a residual and the figures in round brackets are the estimated standard errors.

(i) It is claimed that households spend approximately one fifth of their income on food. Test this claim using the data from households at the 5% significance level.

(You may assume the relevant critical value is 2.0).

At this point, you may be wondering what all these theoretical derivations are useful for. This brief example aims to clarify that. For the aspiring econometrician, the theory behind hypothesis testing may be wonderful. However, for all of you who may consider to do empirical work, understanding why these results are important and what are they used for is crucial.

Our previous result has suggested us that the variable $\frac{(\hat{\beta}-\beta)}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}}}$ is going to let us test whether $\hat{\beta}$ is substantially different (statistically speaking) from β. As economists, the relevant test is most widely used to analyse whether a given variable has any effect at all in our dependent variable. For instance, one may wonder whether years of education (x_t) has any effect on earnings (y_t). For that, the usual practice in the profession is to grab datasets and run an ols (or a variant of) regression between both x_t and y_t, which are observable. Within this setup, the test on whether $\hat{\beta}$ is different from 0 allows us to test whether an extra year of education influences earnings. Notice that we are assuming, ex-ante, that $\beta = 0$. This is normally done as we normally try to unravel new knowledge and, hence, $\beta = 0$ serves as a useful benchmark: Given that the previous knowledge in the literature is that this variable has no effect, I use some new data to test whether that statement is statistically correct or false.

In the proposed case, $\frac{1}{5} = 0.2$. Hence, the ex-ante claimed value of β is $\beta = 0.2$. Our test, then, would involve comparing the estimated value against the ex-ante hypothesized one. More formally, we write the null hypothesis as:

$$H_0\colon \beta = 0.2$$

The alternative hypothesis is that the previous relationship does not hold:

$$H_1\colon \beta \neq 0.2$$

We normally say that we test the null hypothesis (that $\beta = 0.2$) against the alternative (that $\beta \neq 0.2$). To do that, we just substitute $\beta = 0.2$ and $\hat{\beta} = 0.25$ (the coefficient from the regression) in the test statistic outlined in the previous section to get:

$$T = \frac{(0.25 - 0.20)}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}$$

Now, notice that we already found that $V[\hat{\beta}] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}$, and that an unbiased estimator of that variance[10] is:

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

Hence, it follows that the standard deviation of $\hat{\beta}$ is:

$$\widehat{SE}[\hat{\beta}] = \sqrt{\hat{V}[\hat{\beta}]} = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}$$

Which is the denominator of the expression above. Hence, the figure in brackets below the coefficient $\hat{\beta}$ is $\widehat{SE}[\hat{\beta}]$. Plugging it in the expression of the test statistic, we get:

$$T = \frac{(0.25 - 0.20)}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}} = \frac{(0.25 - 0.20)}{\widehat{SE}[\hat{\beta}]} = \frac{0.05}{0.05} = 1$$

[10] You will study this on later courses. At the moment, it suffices to know that $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(\hat{\varepsilon}_t)^2}{n-2}$, and that $E[\hat{\sigma}^2] = E\left[\frac{\sum_{i=1}^{n}(\hat{\varepsilon}_t)^2}{n-2}\right] = \sigma^2$

As $1 < 2$, this implies that our test statistic is lower than the critical value. Hence, we cannot reject the null hypothesis that $\hat{\beta} = 0.20$. In other words, the statistical evidence is not able to refute the ex-ante knowledge we had about the relationship between household income and household expenditure on food.

(ii) Detail how the coefficient of determination R^2 can be used to test whether the variables Y and X are related. Test the hypothesis of no relationship at the 5% significance level.

The relationship between to variables is normally measured by the correlation coefficient:

$$r = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{n}(x_t - \bar{x})^2 \sum_{t=1}^{n}(y_t - \bar{y})^2}}$$

In order to see how the coefficient of determination can be used to test whether variables Y and X are related, we need to show several statements.

First, that the coefficient of determination is equal to the square of the coefficient of correlation:

$$r^2 = R^2$$

Second, that the coefficient of determination determines a test statistic that can be used instead of the one covered in part (c), the F statistic. The F statistic serves the same purpose as the T statistic and it is defined as the square of the T statistic:

$$F = T^2 = \left(\frac{(\hat{\beta} - \beta)}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}} \right)^2 = \frac{R^2}{1 - R^2}$$

All in all, the next sections will i) show the direct relationship between the coefficient of correlation and the coefficient of determination, ii) show that the square of the T statistic is the F statistic, iii) showing the direct relationship between the F statistic and the coefficient of determination and iv) some concluding comments on how to use the F statistic to make analogous hypothesis tests as the one covered in section (d) i).

a. The relationship between the coefficient of determination and the coefficient of determination

We start by recalling the formula for the coefficient of determination:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

By definition, we know that:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t$$

By the optimisation process of OLS, we know that:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{i=1}^{n}(x_t - \bar{x})^2}$$

Substituting the formula for $\hat{y}_t$ in the numerator of R^2, we get:

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{\alpha} + \hat{\beta}x_t - \bar{y}\right)^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

Substituting the OLS formula for $\hat{\alpha}$ in the numerator of R^2, we get:

$$R^2 = \frac{\sum_{i=1}^{n}\left(\bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x_t - \bar{y}\right)^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

Noticing that the $\bar{y}$'s cancel out, the formula collapses to:

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{\beta}x_t - \hat{\beta}\bar{x}\right)^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

Using $\hat{\beta}$ as a common factor, we get:

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{\beta}(x_t - \bar{x})\right)^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

Using the rules of calculus[11], we get:

$$R^2 = \frac{\sum_{i=1}^{n}\hat{\beta}^2(x_t - \bar{x})^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

As $\hat{\beta}^2$, we can take it out of the summation operator to get:

$$R^2 = \frac{\hat{\beta}^2\sum_{i=1}^{n}(x_t - \bar{x})^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

[11] Remember that $(ax)^2 = a^2x^2$

Substituting the formula for the OLS estimator $\hat{\beta}$ in the previous expression yields:

$$R^2 = \frac{\left(\frac{\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{i=1}^{n}(x_t - \bar{x})^2}\right)^2 \sum_{i=1}^{n}(x_t - \bar{x})^2}{\sum_{i=1}^{n}(y_t - y)^2}$$

Which is equivalent to:

$$R^2 = \frac{\frac{(\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y}))^2}{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^2} \sum_{i=1}^{n}(x_t - \bar{x})^2}{\sum_{i=1}^{n}(y_t - y)^2} = \frac{(\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y}))^2 \frac{\sum_{i=1}^{n}(x_t - \bar{x})^2}{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^2}}{\sum_{i=1}^{n}(y_t - y)^2}$$

Further simplifying gets us to:

$$R^2 = \frac{(\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y}))^2 \frac{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^1}{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^2}}{\sum_{i=1}^{n}(y_t - y)^2} = \frac{\frac{(\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y}))^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}{\sum_{i=1}^{n}(y_t - y)^2}$$

Using the rules of calculus, we can rewrite the previous expression as:

$$R^2 = \frac{(\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y}))^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2 \sum_{i=1}^{n}(y_t - y)^2}$$

Taking the square root of the previous expression yields:

$$\sqrt{R^2} = \frac{\sum_{i=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_t - \bar{x})^2 \sum_{i=1}^{n}(y_t - y)^2}} = r$$

Hence, if $r = \sqrt{R^2}$, then squaring both sides yields:

$$R^2 = r^2$$

Hence, one way in which the coefficient of determination can be used to test whether two variables X and Y are related is because the coefficient of determination is nothing but the square of the coefficient of correlation! Hence, it will always provide a measure on how well two variables are related.

b. The relationship between the F statistic and the T statistic

However, note that the previous section only highlights that the coefficient of determination is influenced by the coefficient of correlation. Still we haven't provided a proper statistical test.

Recall the classical linear regression model:

$$y_t = \alpha + \beta x + \varepsilon_t$$

In very simple terms, analysing whether y_t and x_t are related is just a matter of determining the sign of the partial derivative of y_t with respect to x_t. Given that

$$\frac{\partial y_t}{\partial x_t} \neq 0$$

Would serve as evidence supporting a relationship between y_t and x_t. Note that, when computing such partial derivative, we get:

$$\frac{\partial y_t}{\partial x_t} = \frac{\partial(\alpha + \beta x + \varepsilon_t)}{\partial x_t} = \beta$$

Which is just the coefficient of x_t. Hence, our statistical test to check whether y_t and x_t are related has to involve a test that analyses whether the coefficient of β is different from 0. Notice that this test is pretty similar to the one proposed before. More generally, the test we used in subsection i) was testing the null hypothesis:

$$H_0: \beta = \beta_0$$

Against the alternative hypothesis:

$$H_1: \beta \neq \beta_0$$

In the previous section, we set $\beta_0 = 0.2$. Here, we would just need to set $\beta_0 = 0$. The test statistic, hence, should be modified accordingly. Generally, the test statistic reads:

$$T = \frac{\left(\hat{\beta} - \beta_0\right)}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}$$

As we set $\beta_0 = 0$, the specific test of hypothesis when testing whether y_t and x_t are related would be:

$$T = \frac{\hat{\beta}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}}$$

This looks fine as far as it goes: we found a test that will help us to analyse whether y_t and x_t are related. However, e haven't showed whether there is a relationship between the coefficient of determination and such a test. The remainder of this section will show that the squared of the proposed test statistic is an F test. The next section will show that the F statistic is indeed related with the coefficient of determination.

We know that an F test takes the form:

$$F = \frac{q}{m} \frac{TSS - RSS}{RSS}$$

Where q and m are, respectively, the degrees of freedom of the denominator and the numerator. Squaring the expression for T yields:

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

Rearranging yields:

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

As defined before, $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}{n-2}$, substituting $\hat{\sigma}^2$ in the previous expression gives us:

$$T^2 = \frac{\hat{\beta}^2 \sum_{i=1}^{n} (x_t - \bar{x})^2}{\frac{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}{n-2}}$$

Rearranging, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right) \frac{\hat{\beta}^2 \sum_{i=1}^{n} (x_t - \bar{x})^2}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Substituting $\hat{\beta}$ by its OLS expression yields:

$$T^2 = \left(\frac{(n-2)}{1}\right) \frac{\left(\frac{\sum_{i=1}^{n} x_t (y_t - \bar{y})}{\sum_{i=1}^{n} (x_t - \bar{x})^2}\right)^2 \sum_{i=1}^{n} (x_t - \bar{x})^2}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Which is equal to:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\frac{(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2 (\sum_{i=1}^{n}(x_t - \bar{x})^2)^1}{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^2}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Which, after simplifying, collapses to:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\frac{(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Noticing that $\frac{(\sum_{i=1}^{n} x_t(y_t-\bar{y}))^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2} = -\frac{(\sum_{i=1}^{n} x_t(y_t-\bar{y}))^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2} + 2\frac{(\sum_{i=1}^{n} x_t(y_t-\bar{y}))^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2}$, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{-\frac{(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2} + 2\frac{(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2}{\sum_{i=1}^{n}(x_t - \bar{x})^2}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Multiplying the first element of the numerator by $\frac{\sum_{i=1}^{n}(x_t-\bar{x})^2}{\sum_{i=1}^{n}(x_t-\bar{x})^2} = 1$ and noticing that $(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2 = \sum_{i=1}^{n} x_t(y_t - \bar{y}) * \sum_{i=1}^{n} x_t(y_t - \bar{y})$, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{-\frac{(\sum_{i=1}^{n} x_t(y_t - \bar{y}))^2}{(\sum_{i=1}^{n}(x_t - \bar{x})^2)^2}\sum_{i=1}^{n}(x_t - \bar{x})^2 + 2\frac{\sum_{i=1}^{n} x_t(y_t - \bar{y})}{\sum_{i=1}^{n}(x_t - \bar{x})^2}\sum_{i=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Substituting $\frac{\sum_{i=1}^{n} x_t(y_t-\bar{y})}{\sum_{i=1}^{n}(x_t-\bar{x})^2}$ by $\hat{\beta}$, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{-\hat{\beta}^2 \sum_{i=1}^{n}(x_t - \bar{x})^2 + 2\hat{\beta}\sum_{i=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Substituting $\sum_{i=1}^{n}(y_t - \bar{y})^2 - \sum_{i=1}^{n}(y_t - \bar{y})^2 = 0$ in the numerator, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t - \bar{y})^2 - \sum_{i=1}^{n}(y_t - \bar{y})^2 - \hat{\beta}^2 \sum_{i=1}^{n}(x_t - \bar{x})^2 + 2\hat{\beta}\sum_{i=1}^{n} x_t(y_t - \bar{y})}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Noticing that:

$$-\sum_{i=1}^{n}(y_t - \bar{y})^2 - \hat{\beta}^2\sum_{i=1}^{n}(x_t - \bar{x})^2 + 2\hat{\beta}\sum_{i=1}^{n} x_t(y_t - \bar{y}) = -\left[\sum_{i=1}^{n}(y_t - \bar{y})^2 \hat{\beta}^2 \sum_{i=1}^{n}(x_t - \bar{x})^2 - 2\hat{\beta}\sum_{i=1}^{n} x_t(y_t - \bar{y})\right]$$

T^2 is equal to:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t - \bar{y})^2 - \left[\sum_{i=1}^{n}(y_t - \bar{y})^2 + \hat{\beta}^2\sum_{i=1}^{n}(x_t - \bar{x})^2 - 2\hat{\beta}\sum_{i=1}^{n} x_t(y_t - \bar{y})\ \right]}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

Including the $\hat{\beta}$'s within the summation terms, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\sum_{i=1}^{n}(y_t-\bar{y})^2 + \sum_{i=1}^{n}\hat{\beta}^2(x_t-\bar{x})^2 - 2\sum_{i=1}^{n}\hat{\beta}x_t(y_t-\bar{y})\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Given that

$$\sum_{i=1}^{n}x_t(y_t-\bar{y}) = \sum_{i=1}^{n}(x_t-\bar{x})(y_t-\bar{y})$$

The previous expression becomes:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\sum_{i=1}^{n}(y_t-\bar{y})^2 + \sum_{i=1}^{n}\hat{\beta}^2(x_t-\bar{x})^2 - 2\sum_{i=1}^{n}\hat{\beta}(x_t-\bar{x})(y_t-\bar{y})\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Noticing that $(a-b)^2 \equiv a^2+b^2-2ab$, we can rewrite the previous expression as:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\left(\sum_{i=1}^{n}(y_t-\bar{y}) - \sum_{i=1}^{n}\hat{\beta}(x_t-\bar{x})\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Using the rules of the summation operator, we can divide the summation terms within the brackets into a single summation term:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\left(\sum_{i=1}^{n}\left((y_t-\bar{y}) - \hat{\beta}(x_t-\bar{x})\right)\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Expanding $\hat{\beta}(x_t-\bar{x})$, we get:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\left(\sum_{i=1}^{n}(y_t-\bar{y}-\hat{\beta}x_t+\widehat{\beta\bar{x}})\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Which is equivalent to:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\left(\sum_{i=1}^{n}(y_t-(\bar{y}-\hat{\beta}\bar{x})-\hat{\beta}x_t)\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Noticing that $\hat{\alpha} = \bar{y} - \widehat{\beta}\bar{x}$, we can rewrite the previous expression as:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t-\bar{y})^2 - \left[\left(\sum_{i=1}^{n}(y_t-\widehat{\alpha}-\hat{\beta}x_t)\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Which can be rewritten as:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t - \bar{y})^2 - \left[\left(\sum_{i=1}^{n}\left(y_t - \left(\hat{\alpha} + \hat{\beta}x_t\right)\right)\right)^2\right]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Noticing that $\hat{\alpha} + \hat{\beta}x_t = \hat{y}_t$, we can rewrite the previous expression as:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t - \bar{y})^2 - [(\sum_{i=1}^{n}(y_t - \hat{y}_t))^2]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Noticing that $y_t - \hat{y}_t = \hat{\varepsilon}_t$, we can rewrite the previous expression as:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{\sum_{i=1}^{n}(y_t - \bar{y})^2 - [(\sum_{i=1}^{n}(\hat{\varepsilon}_t))^2]}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

Noticing that $TSS = \sum_{i=1}^{n}(y_t - \bar{y})^2$ and $RSS = \sum_{t=1}^{n}\hat{\varepsilon}_t^2$, the previous expression becomes:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{TSS - RSS}{RSS} \equiv F$$

Also, note that, in its general form, a t-student variable is formed in the following way:

$$T = \frac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}}$$

Squaring both sides gives:

$$T^2 = \frac{\left(N(0,1)\right)^2}{\left(\sqrt{\frac{\chi^2_{(n)}}{n}}\right)^2} = \frac{\chi^2_{(1)}}{\frac{\chi^2_{(n)}}{n}}$$

As the numerator only includes one normal (0,1) squared. As $\chi^2_{(1)} = \frac{\chi^2_{(1)}}{1}$, the previous equation is equivalent to:

$$T^2 = \frac{\left(N(0,1)\right)^2}{\left(\sqrt{\frac{\chi^2_{(n)}}{n}}\right)^2} = \frac{\frac{\chi^2_{(1)}}{1}}{\frac{\chi^2_{(n)}}{n}} = \left(\frac{n}{1}\right)\left(\frac{\chi^2_{(1)}}{\chi^2_{(n)}}\right)$$

As a fraction of two chi square variables multiplied by the ratio of their degrees of freedom is the definition of a variable following an F distribution, then it follows that:

$$T^2 = \left(\frac{(n-2)}{1}\right)\frac{TSS - RSS}{RSS} \equiv F = \left(\frac{n}{1}\right)\left(\frac{\chi^2_{(1)}}{\chi^2_{(n)}}\right) \sim F_{(1,n-2)}$$

Hence, in this section we have demonstrated that the square of the T statistic is the F statistic and that the F statistic follows a F distribution.

c. The relationship between the F statistic and the coefficient of determination

This last section demonstrates that the F statistic is directly related with the coefficient of determination. More specifically,

$$F = \frac{R^2}{1 - R^2}$$

Recalling the F statistic:

$$F = \left(\frac{(n-2)}{1}\right)\frac{TSS - RSS}{RSS}$$

Multiplying the RSS in the denominator by $\frac{TSS}{TSS} = 1$, we get:

$$F = \left(\frac{(n-2)}{1}\right)\frac{TSS - RSS * \frac{TSS}{TSS}}{RSS}$$

Using TSS as a common factor, we get:

$$F = \left(\frac{(n-2)}{1}\right)\frac{TSS * \left(1 - \frac{RSS}{TSS}\right)}{RSS}$$

By noticing that $R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$, we get:

$$F = \left(\frac{(n-2)}{1}\right)\frac{TSS * (R^2)}{RSS}$$

Using the rules of calculus, the previous equation is equivalent to:

$$F = \left(\frac{(n-2)}{1}\right)\frac{R^2}{\frac{RSS}{TSS}}$$

As $R^2 = 1 - \frac{RSS}{TSS}$, then it follows that $1 - R^2 = \frac{RSS}{TSS}$. Hence,

$$F = \left(\frac{(n-2)}{1}\right)\frac{R^2}{1-R^2} = \left(\frac{(n-2)}{1}\right)\frac{r^2}{1-r^2}$$

d. Concluding remarks for hypothesis testing

In this exercise, we have learned several things. First, that the coefficient of determination and the coefficient of correlation are related. Second, that to test whether two variables are related we can use a T test or, alternatively, an F test. Third, that the F test that serves to analyse whether to variables are related statistically is, indeed, related with the coefficient of correlation.

Hence, in order to test whether two variables are related the coefficient of correlation is a key element to take into account. Not only provides a measurement of correlation of two variables, but can be used, as shown in the equation above, to provide a statistical test of the relation between two variables.

Finally, the coefficient of determination also has a dual role. It provides a measurement on how much part of the variance of a variable is to be explained by the variance of another variable. And, as well, it provides a statistical test to analyse the relation between two variables.

At the 5% level, we have:

$$T^2 = \left(\frac{0.25}{0.05}\right)^2 = (5)^2 = 25 = F$$

Also, the critical value is the square of the critical value for the t distribution analysed before (4).

Hence, as $25 > 4$, it follows that we can reject the null hypothesis of no correlation in favour of the alternative hypothesis of correlation.