PROPOSED SOLUTIONS
HOMEWORK 4
GRA6039 ECONOMETRICS WITH PROGRAMMING
AUTUMN 2020

EMIL A. STOLTENBERG

Solutions to Ex. 1. Let $X_1, \dots, X_n$ be random variables, numbers, observations.
(a) Let's try with $n = 4$, then

$$\sum_{i=1}^{3}(X_{i+1} - X_i) = X_2 - X_1 + X_3 - X_2 + X_4 - X_3 = X_4 - X_1.$$

(b) Let $a_i = \sum_{j=i}^{3} X_j$ for $i = 1, 2, 3$. Then

$$\begin{aligned}\sum_{i=1}^{3} a_i = a_1 + a_2 + a_3 &= \sum_{j=1}^{3} X_j + \sum_{j=2}^{3} X_j + \sum_{j=3}^{3} X_j \\ &= (X_1 + X_2 + X_3) + (X_2 + X_3) + X_3 = X_1 + 2X_2 + 3X_3.\end{aligned}$$

(c) Generalise what you found in (b). Or

$$\begin{aligned}\sum_{i=1}^{n} iX_i &= 1X_1 + 2X_2 + 3X_3 + 4X_4 \cdots + nX_n \\ &= \sum_{i=1}^{n} X_i + \{X_2 + 2X_3 + 3X_4 \cdots + (n-1)X_n\} \\ &= \sum_{i=1}^{n} X_i + \sum_{i=2}^{n} X_i + \{X_3 + 2X_4 \cdots + (n-2)X_n\} \\ &= \sum_{i=1}^{n} X_i + \sum_{i=2}^{n} X_i + \cdots + + \sum_{i=n-1}^{n} X_i + \sum_{i=n}^{n} X_i \\ &= \sum_{j=1}^{n} \sum_{i=j}^{n} X_i.\end{aligned}$$

Date: October 5, 2020.

Solutions to Ex. 2. Let $X_1, \dots, X_n$ an $Y_1, \dots, Y_m$ be random variables, and define the random variables $Z_1, \dots, Z_{n+m}$ as follows,

$$Z_1 = X_1, \dots, Z_n = X_n,\ Z_{n+1} = Y_1, \dots, Z_{n+m} = Y_m.$$

(a)

$$\begin{aligned}\bar{Z}_{n+m} &= \frac{1}{n+m}\sum_{i=1}^{n+m} Z_i = \frac{1}{n+m}(\sum_{i=1}^{n} X_i + \sum_{i=1}^{m} Y_i) \\ &= \frac{1}{n+m}(n\bar{X}_n + m\bar{Y}_m) = \frac{n}{n+m}\bar{X}_n + \frac{m}{n+m}\bar{Y}_m.\end{aligned}$$

in terms of $\bar{X}_n$ and $\bar{Y}_m$. **(b)** When $n = m$,

$$\frac{1}{2}(\bar{X}_n + \bar{Y}_m) = \bar{Z}_{n+m}.$$

(c) Let a be some constant, then

$$\begin{aligned}\sum_{i=1}^{n}(X_i - a)^2 &= \sum_{i=1}^{n}(X_i - \bar{X}_n + \bar{X}_n - a)^2 \\ &= \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + 2(\bar{X}_n - a)\sum_{i=1}^{n}(X_i - \bar{X}_n) + n(\bar{X}_n - a)^2 \\ &= (n-1)s_X^2 + n(\bar{X}_n - a)^2,\end{aligned}$$

because $\sum_{i=1}^{n}(X_i - \bar{X}_n) = 0$ and $(n-1)s_X^2 = \sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.
(d) Look at

$$(n+m-1)s_Z^2 = \sum_{i=1}^{n+m}(Z_i - \bar{Z}_{n+m})^2 = \sum_{i=1}^{n}(X_i - \bar{Z}_{n+m})^2 + \sum_{i=1}^{m}(Y_i - \bar{Z}_{n+m})^2.$$

It suffices to only look at one of the sums on the right. Use what we found in (c), with $\bar{Z}_{n+m}$ playing the role of a,

$$\begin{aligned}\sum_{i=1}^{n}(X_i - \bar{Z}_{n+m})^2 &= \sum_{i=1}^{n}(X_i - \bar{X}_n + \bar{X}_n - \bar{Z}_{n+m})^2 = (n-1)s_X^2 + n(\bar{X}_n - \bar{Z}_{n+m})^2 \\ &= (n-1)s_X^2 + n\big(\bar{X}_n - \frac{n}{n+m}\bar{X}_n + \frac{m}{n+m}\bar{Y}_m\big)^2 \\ &= (n-1)s_X^2 + \frac{nm^2}{(n+m)^2}(\bar{X}_n - \bar{Y}_m)^2,\end{aligned}$$

from which we see that

$$\sum_{i=1}^{m}(Y_i - \bar{Z}_{n+m})^2 = (m-1)s_Y^2 + \frac{mn^2}{(n+m)^2}(\bar{X}_n - \bar{Y}_m)^2.$$

Some algebra, e.g. $nm^2 + mn^2 = nm(n+m)$, then gives,

$$(n+m-1)s_Z^2 = (n-1)s_X^2 + (m-1)s_Y^2 + \frac{nm}{(n+m)}(\bar{X}_n - \bar{Y}_m)^2.$$

(e) Run and understand the Matlab code.

Solutions to Ex. 3. Suppose you have a coin whose probability of showing heads equals θ (some unknown parameter). We represent one toss of this coin by the random variable

$$X = \begin{cases} 0, & \text{if tails,} \\ 1, & \text{if heads,} \end{cases}$$

which means that

$$\Pr(X = 1) = \theta.$$

We decide to toss this coin until we get a heads up, then stop. By so deciding, we can define a new random variable,

$$Y = \text{the numbers of tosses until we get heads up,}$$

so that Y takes its values in $\{1, 2, 3, \ldots\}$. For example, if we toss tails, tails, heads, then $Y = 3$.

(a) We tacitly understand that the tosses are independent, and we can represent the ith toss by the rv X_i, so that $\Pr(X_i = 1) = \theta$. The few first

$$\begin{aligned}
\Pr(Y = 1) &= \Pr(X_1 = 1) = \theta, \\
\Pr(Y = 2) &= \Pr(X_1 = 0)\Pr(X_2 = 1) = (1 - \theta)\theta, \\
\Pr(Y = 3) &= \Pr(X_1 = 0)\Pr(X_2 = 0)\Pr(X_3 = 1) = (1 - \theta)^2\theta, \\
\Pr(Y = 4) &= \Pr(X_1 = 0)\Pr(X_2 = 0)\Pr(X_3 = 0)\Pr(X_4 = 1) = (1 - \theta)^3\theta,
\end{aligned}$$

(b) from which we see a pattern, namely that

$$\Pr(Y = y) = (1 - \theta)^{y-1}\theta.$$

The pmf of Y is then

$$f_\theta(y) = (1 - \theta)^{y-1}\theta, \quad \text{for } y = 1, 2, 3, \ldots,$$

and $f_\theta(y) = 0$ when y does not equal $1, 2, 3, \ldots$.

(c) We know that

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}, \quad \text{and} \quad \sum_{k=0}^{\infty} x^k = \frac{1}{1 - x},$$

provided $x \neq 1$ and $|x| < 1$, respectively. To show that $f_\theta(y)$ is a pmf we must show that $f_\theta(y) \geq 0$ for all y, and that is sums to one. Since $0 \leq \theta \leq 1$, $f_\theta(y)$ is non-negative. For the second,

$$\begin{aligned}
\sum_{y=1}^{\infty}(1 - \theta)^{y-1}\theta &= \frac{\theta}{1 - \theta}\sum_{y=1}^{\infty}(1 - \theta)^y = \frac{\theta}{1 - \theta}\Big\{\sum_{y=0}^{\infty}(1 - \theta)^y - 1\Big\} \\
&= \frac{\theta}{1 - \theta}\Big\{\frac{1}{\theta} - 1\Big\} = \frac{\theta}{1 - \theta}\frac{1 - \theta}{\theta} = 1.
\end{aligned}$$

(d) Here we show that $\mathrm{E}\,Y = \sum_{y=1}^{\infty} y f(y) = 1/\theta$. It is important for what follows that since $0 < \theta < 1$, then $0 < 1 - \theta < 1$.

$$\mathrm{E}\,[Y] = \sum_{y=1}^{\infty} y f(y) = \sum_{y=1}^{\infty} y(1 - \theta)^{y-1}\theta = \frac{\theta}{1 - \theta}\sum_{y=1}^{\infty} y(1 - \theta)^y,$$

Let us therefore look at $\sum_{y=1}^{\infty} y(1-\theta)^y$. For this sum we'll use the result from Ex. 1(c), generalised to infinite sums,

$$\begin{aligned}\sum_{y=1}^{\infty} y(1-\theta)^y &= \sum_{k=1}^{\infty}\sum_{y=k}^{\infty}(1-\theta)^y = \sum_{k=1}^{\infty}\Big\{\sum_{y=1}^{\infty}(1-\theta)^y - \sum_{y=1}^{k-1}(1-\theta)^y\Big\} \\ &= \sum_{k=1}^{\infty}\Big\{\sum_{y=0}^{\infty}(1-\theta)^y - \sum_{y=0}^{k-1}(1-\theta)^y\Big\} \\ &= \sum_{k=1}^{\infty}\Big\{\frac{1}{\theta} - \frac{1-(1-\theta)^k}{\theta}\Big\} = \sum_{k=1}^{\infty}\frac{(1-\theta)^k}{\theta} = \frac{1}{\theta}\sum_{k=1}^{\infty}(1-\theta)^k \\ &= \frac{1}{\theta}\Big\{\sum_{k=0}^{\infty}(1-\theta)^k - 1\Big\} = \frac{1}{\theta}\Big\{\frac{1}{\theta} - 1\Big\} = \frac{1-\theta}{\theta^2}.\end{aligned}$$

This shows that

$$\frac{1-\theta}{\theta}\mathrm{E}\,[Y] = \frac{1-\theta}{\theta^2},$$

and therefore $\mathrm{E}\,[Y] = 1/\theta$.

(e) We have independent $Y_1, \ldots, Y_n$ from $f_\theta(y)$. First

$$\log f_\theta(y) = \log\{(1-\theta)^{y-1}\theta\} = (y-1)\log(1-\theta) + \log\theta,$$

and the log-likelihood function is

$$\ell_n(\theta) = \sum_{i=1}^{n}\log f_\theta(Y_i) = \log(1-\theta)\sum_{i=1}^{n}(Y_i - 1) + n\log\theta = \log(1-\theta)n(\bar{Y}_n - 1) + n\log\theta.$$

(f) Find the first derivative of $\ell_n(\theta)$, set it equal to zero,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_n(\theta) = -\frac{n(\bar{Y}_n - 1)}{1-\theta} + \frac{n}{\theta} = 0.$$

Solve for θ to find the MLE, it is $\widehat{\theta}_n = 1/\bar{Y}_n$.

(g) Show that $\widehat{\theta}_n \to_p \theta$, i.e. that $\widehat{\theta}_n$ is consistent for θ. Note first that

$$\mathrm{Var}(\bar{Y}_n) = \frac{1-\theta}{n\theta^2},$$

which is finite, so the Law of large numbers (LLN) applies. Can argue in two ways: (1) $\bar{Y}_n \to_p 1/\theta$ by the LLN, and $g(x) = 1/x$ is a continuous function (except at $x = 0$). We know that if $X_n \to_p a$, and $h(x)$ is a continuous function, then $h(X_n) \to_p h(a)$ (see notes from Lecture 5, and Wooldridge (2019, Property PLIM.1, p. 722)). Thus,

$$\widehat{\theta}_n = g(\bar{Y}_n) \overset{p}{\to} g(1/\theta) = \frac{1}{1/\theta} = \theta.$$

If we did not know about Property PLIM.1, but only knew Chebyshev's inequality as presented in Lecture 4 (Lemma 4.2 in the machine written lecture notes), we could argue as follows. Since $Y_i \geq 1$ for all i, the empirical mean $\bar{Y}_n \geq 1$. Then,

$$|\widehat{\theta}_n - \theta| = \frac{|1-\theta\bar{Y}_n|}{|\bar{Y}_n|} \leq |1-\theta\bar{Y}_n|,$$

and we must therefore have the following inclusion of events: for any $\varepsilon > 0$,

$$\{|\widehat{\theta}_n - \theta| \geq \varepsilon\} \subset \{|1 - \theta\bar{Y}_n| \geq \varepsilon\},$$

Now, $\mathrm{E}\,[\theta\bar{Y}_n] = 1$ and $\mathrm{Var}(\theta\bar{Y}_n) = (1-\theta)/n$, so

$$\Pr(|\widehat{\theta}_n - \theta| \geq \varepsilon) \leq \Pr(|1 - \theta\bar{Y}_n| \geq \varepsilon) \leq \frac{1-\theta}{\varepsilon^2 n},$$

where the second inequality comes from Chebyshev's inequality. The right hand side tends to zero as $n \to \infty$, which shows that $\widehat{\theta}_n$ is consistent for θ.

Solutions to Ex. 3. Let $Y_1, \ldots, Y_n$ be independent random variables; and let $x_1, \ldots, x_n$ be some numbers, at least one of which does not equal zero. Assume that $Y_i \sim \mathrm{N}(\theta x_i, \sigma^2)$ for $i = 1, \ldots, n$. That is, the density of the ith random variable Y_i is

$$f_i(y; \theta, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{ -\frac{1}{2\sigma^2}(y - \theta x_i)^2 \Big\},$$

where $\sigma > 0$ and $\theta \in \mathbb{R}$. In this exercise we will study the maximum likelihood estimators of θ and σ^2.

(a) The logarithm of the ith density is

$$\log f_i(y; \theta, \sigma^2) = -\frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(y - \theta x_i)^2 - \log\sqrt{2\pi},$$

using that $(1/2)\log\sigma^2 = \log\sigma$. Then

$$\ell_n(\theta, \sigma^2) = \sum_{i=1}^n \log f_i(Y_i; \theta, \sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \theta x_i)^2 - n\log\sqrt{2\pi},$$

and using the chain rule for differentiation, we get

$$\frac{\partial}{\partial\theta}\ell_n(\theta, \sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^n (Y_i - \theta x_i)x_i.$$

The expectation of $\partial\ell(\theta, \sigma^2)/\partial\theta$ is

$$\mathrm{E}\,\frac{\partial}{\partial\theta}\ell_n(\theta, \sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^n (\mathrm{E}\,[Y_i] - \theta x_i)x_i = \frac{1}{\sigma^2}\sum_{i=1}^n (\theta x_i - \theta x_i)x_i = 0.$$

(b) Set $\partial\ell(\theta, \sigma^2)/\partial\theta = 0$ and solve for θ,

$$\frac{1}{\sigma^2}\sum_{i=1}^n Y_i x_i - \theta\frac{1}{\sigma^2}\sum_{i=1}^n x_i^2 = 0,$$

so the MLE is

$$\widehat{\theta}_n = \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2}.$$

Define

$$a_i = \frac{x_i}{\sum_{j=1}^n x_j^2}, \quad \text{for } i = 1, \ldots, n,$$

then

$$\widehat{\theta}_n = \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2} = \sum_{i=1}^n a_i Y_i,$$

and we see that we can use Prop. 2.3 in the Lecture notes,

$$E[\widehat{\theta}_n] = \sum_{i=1}^n a_i \mathrm{E}[Y_i] = \sum_{i=1}^n a_i x_i \theta = \theta \sum_{i=1}^n a_i x_i = \theta \sum_{i=1}^n \frac{x_i^2}{\sum_{j=1}^n x_i^2} = \theta,$$

which shows that $\widehat{\theta}_n$ is unbiased for θ.

(c) Similarly, due to the independence of $Y_1, \dots, Y_n$ (see HW2 Ex. 3(f)),

$$\mathrm{Var}(\widehat{\theta}_n) = \sum_{i=1}^n a_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^n a_i^2 = \sigma^2 \sum_{i=1}^n \frac{x_i^2}{(\sum_{j=1}^n x_j^2)^2} = \frac{\sigma^2}{\sum_{j=1}^n x_j^2}.$$

(d) Suppose that $\sum_{i=1}^n x_i \to \infty$ as $n \to \infty$, then for any $\varepsilon > 0$,

$$\Pr(|\widehat{\theta}_n - \theta| \geq \varepsilon) \leq \frac{\widehat{\theta}_n}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 \sum_{j=1}^n x_j^2} \to 0,$$

as $n \to \infty$. This shows that $\widehat{\theta}_n \to \theta$, i.e. $\widehat{\theta}_n$ is consistent for θ.

(e) Differentiate $\ell_n(\theta, \sigma^2)$ with respect to σ^2,

$$\frac{\partial}{\partial \sigma^2} \ell_n(\theta, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (Y_i - \theta x_i)^2.$$

Setting $\partial \ell_n(\theta, \sigma^2)/\partial \sigma^2 = 0$, solving for σ^2, and inserting the estimator for θ, yields the MLE,

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \widehat{\theta}_n x_i)^2.$$

(f) Recall that $\partial \ell_n(\theta, \sigma^2)/\partial \theta = \sum_{i=1}^n (Y_i - \theta x_i) x_i$ evaluated in $\widehat{\theta}_n$ equals zero, that is $\partial \ell_n(\widehat{\theta}_n, \sigma^2)/\partial \theta = 0$ (some people prefer $\partial \ell_n(\theta, \sigma^2)/\partial \theta\big|_{\theta = \widehat{\theta}_n} = 0$, or the like)

$$\begin{aligned}
\widehat{\sigma}_n^2 &= \frac{1}{n} \sum_{i=1}^n (Y_i - \widehat{\theta}_n x_i)^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \theta x_i + \theta x_i - \widehat{\theta}_n x_i)^2 \\
&= \frac{1}{n} \sum_{i=1}^n \big\{ (Y_i - \theta x_i)^2 + (\widehat{\theta}_n - \theta)^2 x_i^2 - 2(Y_i - \theta x_i)(\widehat{\theta}_n - \theta) x_i \big\} \\
&= \frac{1}{n} \sum_{i=1}^n \big\{ (Y_i - \theta x_i)^2 + (\widehat{\theta}_n - \theta)^2 x_i^2 - 2(Y_i - \widehat{\theta}_n x_i + \widehat{\theta}_n x_i - \theta x_i)(\widehat{\theta}_n - \theta) x_i \big\} \\
&= \frac{1}{n} \sum_{i=1}^n \big\{ (Y_i - \theta x_i)^2 + (\widehat{\theta}_n - \theta)^2 x_i^2 - 2(\widehat{\theta}_n - \theta)^2 x_i^2 \big\} - \frac{2(\widehat{\theta}_n - \theta)}{n} \frac{\partial}{\partial \theta} \ell_n(\widehat{\theta}_n, \sigma^2) \\
&= \frac{1}{n} \sum_{i=1}^n \big\{ (Y_i - \theta x_i)^2 - (\widehat{\theta}_n - \theta)^2 x_i^2 \big\} = \frac{\sigma^2}{n} \Big\{ \sum_{i=1}^n \frac{(Y_i - \theta x_i)^2}{\sigma^2} - \frac{(\widehat{\theta}_n - \theta)^2}{\sigma^2 / \sum_{i=1}^n x_i^2} \Big\} \\
&= \frac{\sigma^2}{n} \Big\{ \sum_{i=1}^n \frac{(Y_i - \theta x_i)^2}{\sigma^2} - \frac{(\widehat{\theta}_n - \theta)^2}{\mathrm{Var}(\widehat{\theta}_n)} \Big\}.
\end{aligned}$$

We can now use Proposition 2.3 in the Lecture notes to find the expectation of $\widehat{\sigma}_n^2$, but first

$$\mathrm{E}\,\frac{(Y_i-\theta x_i)^2}{\sigma^2} = \frac{1}{\sigma^2}\mathrm{E}\,(Y_i-\theta x_i)^2 = \frac{1}{\sigma^2}\mathrm{Var}(Y_i) = 1,$$

and

$$\mathrm{E}\,\frac{(\widehat{\theta}_n-\theta)^2}{\mathrm{Var}(\widehat{\theta}_n)} = \frac{1}{\mathrm{Var}(\widehat{\theta}_n)}\,\mathrm{E}\,(\widehat{\theta}_n-\theta)^2 = \frac{1}{\mathrm{Var}(\widehat{\theta}_n)}\mathrm{Var}(\widehat{\theta}_n) = 1.$$

Then

$$\mathrm{E}\,[\widehat{\sigma}_n^2] = \frac{\sigma^2}{n}\Big(\sum_{i=1}^n \mathrm{E}\,\Big\{\frac{(Y_i-\theta x_i)^2}{\sigma^2}\Big\} - \mathrm{E}\,\Big\{\frac{(\widehat{\theta}_n-\theta)^2}{\mathrm{Var}(\widehat{\theta}_n)}\Big)\Big\} = \frac{\sigma^2}{n}(n-1).$$

(g) Since $\mathrm{E}\,[\widehat{\sigma}_n^2] = (n-1)\sigma^2/n$, we see that $\widehat{\sigma}_n^2$ is a biased estimator for σ^2. To show that $\widehat{\sigma}_n^2 \to_p \sigma^2$ we use Property PLIM.2 in Wooldridge (2019, p. 724). First

$$\widehat{\sigma}_n^2 - \sigma^2 == (\widehat{\sigma}_n^2 - \mathrm{E}\,[\widehat{\sigma}_n^2]) + (\mathrm{E}\,[\widehat{\sigma}_n^2] - \sigma^2) = (\widehat{\sigma}_n^2 - \mathrm{E}\,[\widehat{\sigma}_n^2]) + (\frac{n-1}{n}\sigma^2 - \sigma^2),$$

here $((n-1)/n\sigma^2 - \sigma^2) = -\sigma^2/n$ is a deterministic sequence that tends to zero, so it also tends to zero in probability. Second, using PLIM.2(i), we only need to show that $\widehat{\sigma}_n^2 - \mathrm{E}\,[\widehat{\sigma}_n^2] \to_p 0$: By Chebyshev's inequality, for any $\varepsilon > 0$,

$$\Pr(|\widehat{\sigma}_n^2 - \mathrm{E}\,[\widehat{\sigma}_n^2]| \geq \varepsilon) \leq \frac{\mathrm{Var}(\widehat{\sigma}_n^2)}{\varepsilon^2} = \frac{2(n-1)\sigma^4}{\varepsilon^2 n^2} \to 0,$$

as $n \to \infty$, and we conclude that $\widehat{\sigma}_n^2 \to_p \sigma^2$, in other words $\widehat{\sigma}_n^2$ is *consistent* for σ^2.

(h) The complete Matlab code for (h)–(k) is given below.

(i) The estimates, based on the data in `hw4_data.txt`, are

Parameter	Estimate
θ	4.5896
σ^2	2.1719
$\{\mathrm{Var}(\widehat{\theta}_n)\}^{1/2}$	0.2534

(j) We are told that

$$\frac{\widehat{\theta}_n - \theta}{\mathrm{se}(\widehat{\theta}_n)} \sim \mathrm{N}(0,1),$$

and that $\Pr(-1.96 \leq Z \leq 1.96) = 0.95$ when $Z \sim \mathrm{N}(0,1)$. A 95% confidence interval for θ is found by isolating θ

$$\begin{aligned}\{-1.96 \leq \frac{\widehat{\theta}_n - \theta}{\mathrm{se}(\widehat{\theta}_n)} \leq 1.96\} &= \{-1.96\,\mathrm{se}(\widehat{\theta}_n) \leq \widehat{\theta}_n - \theta \leq 1.96\,\mathrm{se}(\widehat{\theta}_n)\}\\ &= \{-\widehat{\theta}_n - 1.96\,\mathrm{se}(\widehat{\theta}_n) \leq -\theta \leq -\widehat{\theta}_n + 1.96\,\mathrm{se}(\widehat{\theta}_n)\}\\ &= \{\widehat{\theta}_n - 1.96\,\mathrm{se}(\widehat{\theta}_n) \leq \theta \leq \widehat{\theta}_n + 1.96\,\mathrm{se}(\widehat{\theta}_n)\},\end{aligned}$$

so that

$$\Pr\{\widehat{\theta}_n - 1.96\,\mathrm{se}(\widehat{\theta}_n) \leq \theta \leq \widehat{\theta}_n + 1.96\,\mathrm{se}(\widehat{\theta}_n)\} = 0.95.$$

Therefore

$$[\widehat{\theta}_n - 1.96\,\mathrm{se}(\widehat{\theta}_n), \widehat{\theta}_n + 1.96\,\mathrm{se}(\widehat{\theta}_n)],$$

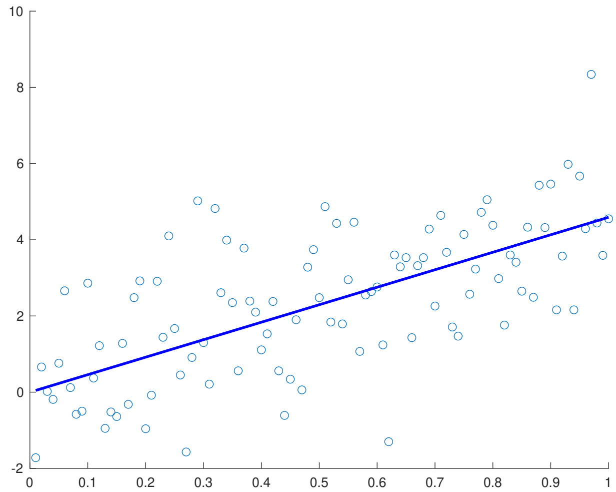

FIGURE 1. A scatter plot of the pairs of (x_i, Y_i) for $i = 1, \ldots, n$ found in `hw4_data.txt`, with the fitted line $\widehat{g}_n(x) = \widehat{\theta}_n x$ overlaid.

is a random interval that will contain θ with 95% probability. It is what is called a 95% confidence interval. A realisation of this interval based on the data in `hw4_data.txt` is

$$[4.09, 5.09].$$

(k) The plot is in Figure 1, and here is the Matlab code

```
cd("~/your_path/");
data = readmatrix("hw4_data.txt");
x = data(:,1);
y = data(:,2);

thetahat = sum(x.*y)/sum(x.^2);
sigma2hat = mean((y - hat.*x).^2);
se_thetahat = sqrt(sigma2hat/sum(x.^2));

% A 95 prct confidence interval
thetahat - 1.96*se_thetahat
thetahat + 1.96*se_thetahat

scatter(x,y)
line(x,thetahat.*x,"Linewidth",2,"Color","b");
saveas(gcf,"~/your_path/hw4scatter.eps","epsc");
```

References

Wooldridge, J. M. (2019). *Introductory Econometrics: A Modern Approach. Seventh Edition.* Cengage Learning, Boston, MA.

Department of Economics, BI Norwegian Business School
Email address: emil.a.stoltenberg@bi.no