PROPOSED SOLUTIONS
HOMEWORK 2
GRA6039 ECONOMETRICS WITH PROGRAMMING
AUTUMN 2020

EMIL A. STOLTENBERG

Solution to Ex. 1. (a) The rv X takes its values in $\mathcal{X} = \{0, 1\}$, with distribution $\Pr(X = 1) = p$. The expectation of X is

$$\mathrm{E}\,X = \sum_{x \in \mathcal{X}} x f(x) = \sum_{x=0}^{1} x f(x) = \sum_{x=0}^{1} x p^x (1-p)^{1-x} = 0 \times (1-p) + 1 \times p = p.$$

(b) The variance of X is

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}\,[(X-p)^2] = \sum_{x=0}^{1} (x-p)^2 p^x (1-p)^{1-x} \\ &= (0-p)^2(1-p) + (1-p)^2 p = p^2(1-p) + (1-p)^2 p \\ &= p(1-p)\{p + (1-p)\} = p(1-p). \end{aligned}$$

(c) We have the function $g(x) = 2x - 1$. With X a rv, then $g(X)$ is a rv. Its expectation is

$$\mathrm{E}\,g(X) = \sum_{x \in \mathcal{X}} g(x) f(x) = \sum_{x=0}^{1} (2x-1) p^x (1-p)^{1-x} = -(1-p) + p = 2p - 1,$$

so when $p = 1/2$, the expectation is $\mathrm{E}\,g(X) = 0$. The variance of $g(X)$ is

$$\begin{aligned} \mathrm{Var}\,g(X) &= \mathrm{E}\,[\{2X - 1 - (2p-1)\}^2] = \mathrm{E}\,[\{2X - 2p\}^2] = \mathrm{E}\,[\{2(X-p)\}^2] \\ &= \sum_{x=0}^{1} 4(x-p)^2 p^x (1-p)^{1-x} = 4 \sum_{x=0}^{1} (x-p)^2 p^x (1-p)^{1-x} \\ &= 4\,\mathrm{Var}\,(X) = 4p(1-p), \end{aligned}$$

so when $p = 1/2$, then $\mathrm{Var}\,g(X) = 1$. We could also do this exercise in a more 'direct' manner using the properties of expectation and variance that we are soon to derive in Ex. 3.

$$\mathrm{E}\,[g(X)] = \mathrm{E}\,[2X - 1] = 2\mathrm{E}\,[X] - 1 = 2p - 1,$$

and

$$\mathrm{Var}\,g(X) = \mathrm{Var}\,(2X - 1) = 4\mathrm{Var}\,(X) = 4p(1-p).$$

Date: September 9, 2020.

Solution to Ex. 2. We are given the function $f(x) = \theta x^{\theta-1}$ for $x \geq 0$, and zero elsewhere, with $\theta > 0$ a parameter. **(a)** To check that $f(x)$ is a pdf we need to verify that $f(x) \geq 0$ for all x, and that $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$. The function $\theta x^{\theta-1} \geq 0$ for all $x \geq 0$ since $\theta > 0$, hence $f(x) \geq 0$ for all x.

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_0^1 \theta x^{\theta-1}\,\mathrm{d}x = \Big|_0^1 x^\theta = 1^\theta - 0^\theta = 1.$$

(b) We see that

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = \begin{cases} 0, & x < 0, \\ \int_0^x \theta y^{\theta-1}\,\mathrm{d}y, & 0 \leq x \leq 1, \\ 1, & x \geq 1, \end{cases} = \begin{cases} 0, & x < 0, \\ x^\theta, & 0 \leq x \leq 1, \\ 1, & x \geq 1. \end{cases}$$

Often, we will just write $F(x) = x^\theta$ for $0 \leq x \leq 1$, with the tacit understanding that since $F(x)$ is a cdf, $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x \geq 1$.

(c) $X \sim F$ means that the rv X has the distribution specified by F.

$$\Pr(X > 1/2) = 1 - \Pr(X \leq 1/2) = 1 - F(1/2) = 1 - (1/2)^\theta.$$

(d) When $X \sim F$,

$$\mathrm{E}\,X = \int_0^1 x\theta x^{\theta-1}\,\mathrm{d}x = \int_0^1 \theta x^\theta\,\mathrm{d}x = \Big|_0^1 \frac{\theta}{\theta+1} x^{\theta+1} = \frac{\theta}{\theta+1}.$$

Let's first find the variance the 'hard way'

$$\begin{aligned}
\mathrm{Var}\,(X) &= \int_0^1 (x - \mathrm{E}\,[X])^2 \theta x^{\theta-1}\,\mathrm{d}x = \int_0^1 \{x^2 - 2x\mathrm{E}\,[X] + (\mathrm{E}\,[X])^2\}\theta x^{\theta-1}\,\mathrm{d}x \\
&= \int_0^1 x^2\theta x^{\theta-1}\,\mathrm{d}x - 2\mathrm{E}\,[X] \int x\theta x^{\theta-1}\,\mathrm{d}x + (\mathrm{E}\,[X])^2 \int_0^1 \theta x^{\theta-1}\,\mathrm{d}x \\
&= \int_0^1 \theta x^{\theta+1}\,\mathrm{d}x - 2(\mathrm{E}\,[X])^2 + (\mathrm{E}\,[X])^2 \\
&= \Big|_0^1 \frac{\theta}{\theta+2} x^{\theta+2} - (\mathrm{E}\,[X])^2 = \frac{\theta}{\theta+2} - \frac{\theta^2}{(\theta+1)^2} = \frac{\theta}{(\theta+2)(\theta+1)^2}.
\end{aligned}$$

The 'easy' is to use what we are to show in Ex. 3 (and basically showed just above), namely that

$$\mathrm{Var}\,(X) = \mathrm{E}\,[X^2] - (\mathrm{E}\,[X])^2,$$

then use that $\mathrm{E}\,[X] = \theta/(\theta+1)$ and compute $\mathrm{E}\,[X^2] = \int_0^1 x^2\theta x^{\theta-1}\,\mathrm{d}x = \theta/(\theta+2)$.

When doing possibly confusing integrals and algebra it is a nice habit to check that what you have done is 'probably' correct by way of simulation. Here is a Matlab script where I, for some value of $\theta > 0$ that I choose, check the expressions for $\mathrm{E}\,[X]$ and $\mathrm{Var}\,(X)$ that we just found.

```
theta = 1.23;
u = rand(100,1); % random uniform rv's on [0,1]
x = u.^(1/theta); % probability integral transform

mean(x); EX = theta/(theta + 1);
```

```
var(x); VarX = theta/((theta + 2)*(theta + 1)^2);

fprintf("%f should be close to %f\n",[mean(x),EX])
fprintf("%f should be close to %f\n",[var(x),VarX])
```

Solution to Ex. 3. Proposition 2.3 in the Lecture notes says that the expectation is linear, that is, for rv's X and Y and constants a, b and c,

$$\mathrm{E}\,[aX + bY + c] = a\,\mathrm{E}\,[X] + b\,\mathrm{E}\,[Y] + c.$$

For what follows, you must remember that $\mathrm{E}\,[X]$ and $\mathrm{E}\,[Y]$ are constants, just like a, b and c above, that is, the expectations are not rv's. We'll use Proposition 2.3 over and over in what follows. **(a)**

$$\begin{aligned}\mathrm{Var}\,(X) &= \mathrm{E}\,[(X - \mathrm{E}\,[X])^2] = \mathrm{E}\,[X^2 - 2\,X\mathrm{E}\,[X] + (\mathrm{E}\,[X])^2] \\ &= \mathrm{E}\,[X^2] - 2\,\mathrm{E}\,[X]\mathrm{E}\,[X] + (\mathrm{E}\,[X])^2 = \mathrm{E}\,[X^2] - (\mathrm{E}\,[X])^2.\end{aligned}$$

(b) For a constant a

$$\begin{aligned}\mathrm{Var}\,(aX) &= \mathrm{E}\,[(aX - \mathrm{E}[aX])^2] = \mathrm{E}\,[(aX - a\mathrm{E}[X])^2] \\ &= \mathrm{E}\,[a^2(X - \mathrm{E}[X])^2] = a^2\mathrm{E}\,[(X - \mathrm{E}\,[X])^2] = a^2\mathrm{Var}\,(X).\end{aligned}$$

(c) For a constant a

$$\begin{aligned}\mathrm{Var}\,(a + X) &= \mathrm{E}\,[(a + X - \mathrm{E}[a + X])^2] = \mathrm{E}\,[(a + X - a - \mathrm{E}[X])^2] \\ &= \mathrm{E}\,[(X - \mathrm{E}[X])^2] = \mathrm{Var}\,(X),\end{aligned}$$

(d)

$$\begin{aligned}\mathrm{Cov}\,(X, Y) &= \mathrm{E}\,\{(X - \mathrm{E}\,[X])(Y - \mathrm{E}\,[Y])\} = \mathrm{E}\,\{XY - X\mathrm{E}\,[Y] - Y\mathrm{E}\,[X] + \mathrm{E}\,[X]\mathrm{E}\,[Y]\} \\ &= \mathrm{E}\,[XY] - \mathrm{E}\,[X]\mathrm{E}\,[Y] - \mathrm{E}\,[Y]\mathrm{E}\,[X] + \mathrm{E}\,[X]\mathrm{E}\,[Y] \\ &= \mathrm{E}\,[XY] - \mathrm{E}\,[X]\mathrm{E}\,[Y].\end{aligned}$$

(e) If X and Y are independent, then $\mathrm{E}\,[XY] = \mathrm{E}\,[X]\mathrm{E}\,[Y]$, and we get

$$\mathrm{Cov}\,(X, Y) = \mathrm{E}\,[XY] - \mathrm{E}\,[X]\mathrm{E}\,[Y] = \mathrm{E}\,[X]\mathrm{E}\,[Y] - \mathrm{E}\,[X]\mathrm{E}\,[Y] = 0.$$

(e)

$$\begin{aligned}\mathrm{Var}\,(aX + bY) &= \mathrm{E}\,[(aX + bY)^2] - (\mathrm{E}\,[aX + bY])^2 \\ &= \mathrm{E}\,[a^2X^2 + b^2Y^2 + 2abXY] - (a\mathrm{E}\,[X] + b\mathrm{E}\,[Y])^2 \\ &= a^2\mathrm{E}\,[X^2] + b^2\mathrm{E}\,[Y^2] + 2ab\mathrm{E}\,[XY] - a^2(\mathrm{E}\,[X])^2 - b^2(\mathrm{E}\,[Y])^2 - 2ab\mathrm{E}\,[X]\mathrm{E}\,[Y] \\ &= a^2\{\mathrm{E}\,[X^2] - (\mathrm{E}\,[X])^2\} + b^2\{\mathrm{E}\,[Y^2] - (\mathrm{E}\,[Y])^2\} + 2ab\{\mathrm{E}\,[XY] - \mathrm{E}\,[X]\mathrm{E}\,[Y]\} \\ &= a^2\,\mathrm{Var}\,(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y).\end{aligned}$$

Importantly, when X and Y are independent,

$$\mathrm{Var}\,(aX + bY) = a^2\mathrm{Var}\,(X) + b^2\mathrm{Var}\,(Y).$$

Solution to Ex. 4. $X_1, \dots, X_n$ are i.i.d. rv's with expectation μ and variance σ^2. As usual, $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$. In this exercise we also use Prop. 2.3 from the Lecture notes, and in particular that $\mathrm{Var}\,(aX + bY) = a^2\mathrm{Var}\,(X) + b^2\mathrm{Var}\,(Y)$ when X and Y are independent, as we established in Ex. 3(e). **(a)**

$$\mathrm{E}\,[\bar{X}_n] = \mathrm{E}\,\big(\frac{1}{n}\sum_{i=1}^n X_i\big) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}\,[X_i] = \frac{1}{n}\sum_{i=1}^n \mu = \frac{n}{n}\mu = \mu.$$

(b)

$$\mathrm{Var}\,(\bar{X}_n) = \mathrm{Var}\,\big(\frac{1}{n}\sum_{i=1}^n X_i\big) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\,(X_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{n}{n^2}\sigma^2 = \frac{\sigma^2}{n}.$$

Notice that here the independence of $X_1, \dots, X_n$ is very important, for we actively use the independence assumption to get the second equality.

(c) Because the variance of the empirical mean, namely σ^2/n becomes smaller and smaller as we increase the sample size n.

Solution to Ex. 5. The rv X has the exponential distribution, that is, its pdf is $f(x) = \theta\exp(-\theta x)$ for $x \geq 0$ and $f(x) = 0$ for $x < 0$, where $\theta > 0$ is some parameter. We will often write $X \sim \mathrm{Expo}(\theta)$.

(a) Clearly, $f(x) \geq 0$ for all x. Moreover,

$$\int_0^\infty \theta\exp(-\theta x)\,\mathrm{d}x = -\Big|_0^\infty \exp(-\theta x) = -0 + 1 = 1.$$

(b) For $x < 0$, $\int_{-\infty}^x f(y)\,\mathrm{d}y = 0$, while for $x > 0$,

$$F(x) = \int_0^x \theta\exp(-\theta y)\,\mathrm{d}y = -\Big|_0^x \exp(-\theta y) = -\exp(-\theta x) + 1 = 1 - \exp(-\theta x).$$

(c) Use integration by parts

$$\begin{aligned}\mathrm{E}\,X &= \int_0^\infty x\theta\exp(-\theta x)\,\mathrm{d}x = -\Big|_0^\infty x\exp(-\theta x) + \int_0^\infty e^{-\theta x}\,\mathrm{d}x \\ &= 0 + \int_0^\infty e^{-\theta x}\,\mathrm{d}x = -\frac{1}{\theta}\Big|_0^\infty e^{-\theta x} = -\frac{1}{\theta}(0-1) = \frac{1}{\theta}.\end{aligned}$$

where we use l'Hôpital's rule to show that,

$$\lim_{x\to\infty} x\exp(-\theta x) = \lim_{x\to\infty}\frac{x}{\exp(\theta x)} = \lim_{x\to\infty}\frac{1}{\theta\exp(x)} = 0.$$

Similarly,

$$\begin{aligned}\mathrm{E}\,X^2 &= \int_0^\infty x^2\theta\exp(-\theta x)\,\mathrm{d}x = -\Big|_0^\infty x^2\exp(-\theta x) + \int_0^\infty 2xe^{-\theta x}\,\mathrm{d}x \\ &= 0 + \int_0^\infty 2xe^{-\theta x}\,\mathrm{d}x = \frac{2}{\theta}\int_0^\infty x\theta e^{-\theta x}\,\mathrm{d}x = \frac{2}{\theta}\,\mathrm{E}\,[X] = \frac{2}{\theta^2}.\end{aligned}$$

because $\lim_{x\to\infty} x^2\exp(-\theta x) = 0$. Using Ex. 3(a),

$$\mathrm{Var}\,(X) = \mathrm{E}\,[X^2] - (\mathrm{E}\,[X])^2 = \frac{2}{\theta^2} - \frac{1}{\theta^2} = \frac{1}{\theta^2}.$$

(d) Define the rv Z by

$$Z = \begin{cases} 1, & \text{if } X \geq \log(2)/\theta, \\ 0, & \text{if } X \geq \log(2)/\theta. \end{cases}$$

Notice that Z is a Bernoulli random variable ('a coin flip') like the random variable we met in Ex. 1. This means that if we can find the success probability $\Pr(Z = 1)$, we can can use Ex. 1 to find $\mathrm{E}\,Z$ and $\mathrm{Var}\,Z$.

$$\begin{aligned} \Pr(Z = 1) &= \Pr(X \geq \log(2)/\theta) = 1 - \Pr(X \leq \log(2)/\theta) = 1 - F(\log(2)/\theta) \\ &= 1 - [1 - \exp\{-\theta(\log(2)/\theta)\}] = \exp\{-\theta(\log(2)/\theta)\} \\ &= \exp(-\log(2)) = \exp(\log(1/2)) = \frac{1}{2}. \end{aligned}$$

Which shows that Z is Bernoulli with success probability 1/2, so from Ex. 1 we get

$$\mathrm{E}\,Z = \frac{1}{2}, \quad \text{and} \quad \mathrm{Var}\,Z = \frac{1}{2}\big(1 - \frac{1}{2}\big) = \frac{1}{4}.$$

Solution to Ex. 6. Let Z_1 and Z_2 be independent standard normal random variables. Set

$$\begin{aligned} X &= \sigma_X Z_1 + \mu_X \\ Y &= \sigma_Y(\rho Z_1 + \sqrt{1 - \rho^2}\, Z_2) + \mu_Y, \end{aligned}$$

where $\sigma_X, \sigma_Y > 0$, the correlation coefficient $\rho \in (-1, 1)$, and μ_X and μ_Y are real numbers.

Since Z_1 and Z_2 are independent independent standard normal random variables,

$$\mathrm{E}\,Z_j = 0, \quad \mathrm{Var}\,(Z_j) = \mathrm{E}\,[Z_j^2] = 1, \quad \text{for j =1,2},$$

and $\mathrm{E}\,[Z_1 Z_2] = \mathrm{E}\,[Z_1]\mathrm{E}\,[Z_2] = 0$.

(a) Use Proposition 2.3 from the Lecture notes

$$\mathrm{E}\,[X] = \mathrm{E}\,[\sigma_X Z_1 + \mu_X] = \sigma_X \mathrm{E}\,[Z_1] + \mu_X = \mu_X,$$

and

$$\mathrm{E}\,[Y] = \mathrm{E}\,[\sigma_Y(\rho Z_1 + \sqrt{1 - \rho^2}\, Z_2) + \mu_Y] = \sigma_Y \rho \mathrm{E}\,[Z_1] + \sigma_Y \sqrt{1 - \rho^2}\, \mathrm{E}\,[Z_2] + \mu_Y = \mu_Y.$$

(b) In this exercise and in (c) the independence of Z_1 and Z_2 is important, and we'll use that $\mathrm{Var}(aZ_1 + bZ_2) = a^2\mathrm{Var}(Z_1) + b^2\mathrm{Var}(Z_2) = a^2 + b^2$, and that $\mathrm{E}\,[Z_1 Z_2] = \mathrm{E}\,[Z_1]\mathrm{E}\,[Z_2] = 0$.

$$\mathrm{Var}\,(X) = \mathrm{Var}\,(\sigma_X Z_1 + \mu_X) = \sigma_X^2 \mathrm{Var}(Z_1) = \sigma_X^2,$$

and

$$\begin{aligned} \mathrm{Var}\,(Y) &= \mathrm{Var}\,(\sigma_Y(\rho Z_1 + \sqrt{1 - \rho^2}\, Z_2) + \mu_Y) \\ &= \sigma_Y^2 \rho^2 \mathrm{Var}(Z_1) + \sigma_Y^2(1 - \rho^2)\mathrm{Var}(Z_2) = \sigma_Y^2 \rho^2 + \sigma_Y^2(1 - \rho^2) = \sigma_Y^2. \end{aligned}$$

(c)

$$\begin{aligned}\operatorname{Cov}(X,Y) &= \mathrm{E}\,[\sigma_X Z_1(\sigma_Y(\rho Z_1 + \sqrt{1-\rho^2}\, Z_2))] \\ &= \sigma_X\sigma_Y\rho \mathrm{E}\,[Z_1^2] + \sigma_X\sigma_Y\sqrt{1-\rho^2}\,\mathrm{E}\,[Z_1 Z_2] \\ &= \rho\sigma_X\sigma_Y + \sigma_X\sigma_Y\sqrt{1-\rho^2}\,\mathrm{E}\,[Z_1]\mathrm{E}\,[Z_2] = \rho\sigma_X\sigma_Y.\end{aligned}$$

We say that (X,Y) has the bivariate normal distribution with expectation vector (μ_X,μ_Y), and covariance matrix

$$\begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix},$$

and write

$$\begin{pmatrix}X \\ Y\end{pmatrix} \sim \mathrm{N}_2\big(\begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}, \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\big).$$

The parameter $-1 \le \rho \le 1$ is the correlation,

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

The probability density function of the bivariate normal distribution is

$$\begin{aligned}f(x,y) = &\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \\ &\times \exp\big\{-\frac{1}{2(1-\rho^2)}\big(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(x-\mu_X)(x-\mu_Y)}{\sigma_X\sigma_Y}\big)\big\}.\end{aligned}$$

Solution to Ex. 7. $X_1,\dots,X_n$ are i.i.d. random variables with the uniform distribution on $[0,\theta]$. The pdf of this distribution is $f(x) = 1/\theta$ for $x \in [0,\theta]$, and $f(x) = 0$ for x outside of $[0,\theta]$.

(a) The cdf of one uniform rv is

$$F(x) = \int_0^x \frac{1}{\theta}\,\mathrm{d}x = \frac{x}{\theta}, \quad \text{for } x \in [0,\theta],$$

and $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x \ge \theta$.

(b) We have the random variable M_n defined to be the largest of the $X_1,\dots,X_n$,

$$M_n = \max_{i\le n} X_i = \max\{X_1,\dots,X_n\}.$$

Now, we want to find the cdf of M_n. Note that if $M_n \le x$ then all the X_is must be smaller than x, so these two events are the same

$$\{M_n \le x\} = \{X_i \le x \text{ for all } i\}.$$

Therefore,

$$\begin{aligned}\Pr(M_n \le x) &= \Pr(X_i \le x \text{ for all } i) = \Pr(X_1 \le x,\dots,X_n \le x) \\ &= \Pr(X_1 \le x)\cdots\Pr(X_n \le x) = F(x)\cdots F(x) = F(x)^n,\end{aligned}$$

where we in the third equality use that $X_1,\dots,X_n$ are independent, and in the fourth equality that they are identically distributed.

Notes on Ex. 8 and Ex. 9. You have to play around in Matlab to learn it. Remember to always save your scripts in an .m-file! Some of the point of Ex. 9 is to see how the empirical mean $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ of an i.i.d. sample $X_1, \ldots, X_n$ centers around the expectation, say $\mu = \mathrm{E}[X_1] = \cdots = \mathrm{E}[X_n]$.

Here is a Matlab script that makes a little 'movie' where we see this. Be careful with copy-pasting Matlab code from .pdf's, it often leads to strange errors (due to 'invisible' symbols). Rather write the code into an .m-file yourself.

```
n_max = 600;
mu = 1.23;
sigma = 3.21;
x = normrnd(mu,sigma,100,1);
sims = 400;

for n= 1:n_max
    x_bar = zeros(1,sims);
    for i = 1:sims
        x = normrnd(mu,sigma,n,1);
        x_bar(i) = mean(x);

    end
    histogram(x_bar,"Normalization","pdf")
    xlim([mu-2.8,mu+2.8]);ylim([0,4]);
    line([mu, mu], [0,4], 'LineWidth', 3, 'Color', 'g');
    pause(0.035)
end
```

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