Homework Assignment 1

Matthew Tiger

September 3, 2015

Problem 1. Let $X \sim \text{Expo}(\lambda)$. Find the mean, $\mathrm{E}(X)$, and the variance, $\mathrm{Var}(X)$, of the random variable X.

Solution. If $X \sim \text{Expo}(\lambda)$, then the probability density function of X, $\text{pdf}(x;\lambda)$, is given by

$$\text{pdf}(x;\lambda) := \begin{cases} \lambda \mathrm{e}^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

By definition,

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \text{pdf}(x;\lambda)\,\mathrm{d}x = \int_0^{\infty} x \left(\lambda \mathrm{e}^{-\lambda x}\right) \mathrm{d}x.$$

Using integration by parts, with $u(x) = x$ and $\mathrm{d}v(x) = \lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x$, we have

$$\int_0^{\infty} x \left(\lambda \mathrm{e}^{-\lambda x}\right) \mathrm{d}x = -x\mathrm{e}^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} -\mathrm{e}^{-\lambda x}\,\mathrm{d}x = -\int_0^{\infty} -\mathrm{e}^{-\lambda x}\,\mathrm{d}x.$$

Evaluating the integral, we derive

$$-\int_0^{\infty} -\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{-\mathrm{e}^{-\lambda x}}{\lambda}\Big|_0^{\infty} = \frac{1}{\lambda}.$$

Therefore,

$$\mathrm{E}(X) = \int_0^{\infty} x \left(\lambda \mathrm{e}^{-\lambda x}\right) \mathrm{d}x = \frac{1}{\lambda}. \tag{1}$$

To calculate $\mathrm{Var}(X)$, we can combine the fact that $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$ with the result in (1) to derive that $\mathrm{Var}(X) = E(X^2) - 1/\lambda^2$. Thus, we need only calculate $\mathrm{E}(X^2)$.

Using the definition of $\mathrm{E}(g(X))$ with $g(X) = X^2$, we have

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 \text{ pdf}(x;\lambda)\,\mathrm{d}x = \int_0^{\infty} x^2 \left(\lambda \mathrm{e}^{-\lambda x}\right) \mathrm{d}x.$$

Again, using integration by parts, with $u(x) = x^2$ and $\mathrm{d}v(x) = \lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x$, we have

$$\int_0^{\infty} x^2 \left(\lambda \mathrm{e}^{-\lambda x}\right) \mathrm{d}x = -x^2\mathrm{e}^{-\lambda x}\Big|_0^{\infty} + 2\int_0^{\infty} x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = 2\int_0^{\infty} x\mathrm{e}^{-\lambda x}\,\mathrm{d}x.$$

Using the result from (1) it is clear that after accounting for the λ constant,

$$\int_0^\infty x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda^2}$$

implying that

$$\mathrm{E}(X^2) = \int_0^\infty x^2\left(\lambda\mathrm{e}^{-\lambda x}\right)\mathrm{d}x = 2\int_0^\infty x\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{2}{\lambda^2}.$$

Therefore,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda}.$$

$\square$

Problem 2. Let $X \sim \mathrm{Bin}(n,p)$. Find the mean, $\mathrm{E}(X)$, and the variance, $\mathrm{Var}(X)$, of the random variable X.

Solution. If $X \sim \mathrm{Bin}(n,p)$, then the probability mass function of X, $\mathrm{pmf}(x;n,p)$, is given by

$$\mathrm{pmf}(x;n,p) := \binom{n}{x}p^x(1-p)^{n-x}.$$

By definition,

$$\mathrm{E}(X) = \sum_{x=0}^{\infty} x\mathrm{pmf}(x;n,p) = \sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x}. \tag{2}$$

Since the first term in the series in (2) is 0, we have

$$\sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=1}^{n} x\binom{n}{x}p^x(1-p)^{n-x}.$$

Rewriting the combination in the above equation in terms of factorials gives

$$\sum_{x=1}^{n} x\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=1}^{n} \frac{xn!}{x!(n-x)!}p^x(1-p)^{n-x}.$$

After cancelling the x term and pulling out an n term from the factorial, we derive

$$\sum_{x=1}^{n} \frac{xn!}{x!(n-x)!}p^x(1-p)^{n-x} = n\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!}p^x(1-p)^{n-x}.$$

The astute observer will notice that

$$\frac{(n-1)!}{(x-1)!(n-x)!} = \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!} = \binom{n-1}{x-1},$$

hence

$$n\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!(n-x)!}p^x(1-p)^{n-x} = n\sum_{x=1}^{n}\binom{n-1}{x-1}p^x(1-p)^{n-x}.$$

We can rewrite the index of the above series as $x+1$ so that

$$n\sum_{x=1}^{n}\binom{n-1}{x-1}p^x(1-p)^{n-x} = n\sum_{x=0}^{n-1}\binom{n-1}{x}p^{x+1}(1-p)^{n-(x+1)}.$$

We can rewrite the above equation as follows

$$n\sum_{x=0}^{n-1}\binom{n-1}{x}p^{x+1}(1-p)^{n-(x+1)} = np\sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{(n-1)-x}.$$

The Binomial Theorem tells us that this series is $(p+(1-p))^{n-1}=1$, hence

$$np\sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{(n-1)-x} = np. \tag{3}$$

Therefore,

$$\mathrm{E}(X) = \sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x} = np. \tag{4}$$

To calculate $\mathrm{Var}(X)$, we can combine the fact that $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$ with the result in (4) to derive that $\mathrm{Var}(X) = E(X^2) - (np)^2$. Thus, we need only calculate $\mathrm{E}(X^2)$.

Using the definition of $\mathrm{E}(g(X))$ with $g(X) = X^2$, we have

$$\mathrm{E}(X^2) = \sum_{x=0}^{\infty} x^2 \ \mathrm{pmf}(x;n,p) = \sum_{x=0}^{n} x^2\binom{n}{x}p^x(1-p)^{n-x}.$$

Using the same techniques to calculate $\mathrm{E}(X)$, i.e. writing out the factorial, cancelling like-terms, regrouping, and changing the index, we can rewrite this series as

$$\begin{aligned}\sum_{x=0}^{n} x^2\binom{n}{x}p^x(1-p)^{n-x} &= \sum_{x=0}^{n} x\cdot x\binom{n}{x}p^x(1-p)^{n-x}\\ &= np\sum_{x=0}^{n-1}(x+1)\binom{n-1}{x}p^x(1-p)^{(n-1)-x}.\end{aligned} \tag{5}$$

Let $y(x)$ be defined as

$$y(x) = \binom{n-1}{x}p^x(1-p)^{(n-1)-x}.$$

Then the linearity of the series in (5) allows us to rewrite as

$$
\begin{aligned}
np\sum_{x=0}^{n-1}(x+1)\binom{n-1}{x}p^x(1-p)^{(n-1)-x} &= np\sum_{x=0}^{n-1}(x+1)y(x) \\
&= np\left(\sum_{x=0}^{n-1}xy(x)+\sum_{x=0}^{n-1}y(x)\right). \qquad (6)
\end{aligned}
$$

Using the same techniques to derive (3), i.e. writing out the factorial, cancelling like-terms, regrouping, changing the index, and using the Binomial Theorem, we can work out that the left sum in (6) is

$$
\begin{aligned}
\sum_{x=0}^{n-1}xy(x) &= \sum_{x=0}^{n-1}x\binom{n-1}{x}p^x(1-p)^{(n-1)-x} \\
&= p(n-1)\sum_{x=0}^{n-2}\binom{n-2}{x}p^x(1-p)^{(n-2)-x} \\
&= p(n-1)(p+1-p)^{n-2} = p(n-1). \qquad (7)
\end{aligned}
$$

Using the Binomial Theorem, we know that the right sum in (6) is

$$
\begin{aligned}
\sum_{x=0}^{n-1}y(x) &= \sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{(n-1)-x} \\
&= (p+1-p)^{n-1} = 1 \qquad (8)
\end{aligned}
$$

Combining the results in (7) and (8) we derive that the sum in (6) is

$$
\begin{aligned}
np\left(\sum_{x=0}^{n-1}xy(x)+\sum_{x=0}^{n-1}y(x)\right) &= np(p(n-1)+1) \\
&= np(np-p+1) \\
&= (np)^2-np^2+np.
\end{aligned}
$$

Therefore,

$$
\mathrm{E}(X^2) = (np)^2-np^2+np.
$$

Combining these results we arrive at

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2)-(\mathrm{E}(X))^2 \\
&= (np)^2-np^2+np-(np)^2 \\
&= -np^2+np \\
&= np(1-p). \qquad (9)
\end{aligned}
$$

□

Problem 3. Let X be a random variable and $c \in \mathbb{R}$. Show that, in the discrete and continuous case for X, $\mathrm{E}(cX) = c\mathrm{E}(X)$ and $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$.

Solution. Suppose first that X is a discrete random variable. Then X has probability mass function $\mathrm{pmf}(x)$ and

$$\mathrm{E}(g(X)) = \sum_{x=0}^{\infty} g(x)\mathrm{pmf}(x). \tag{10}$$

Thus, to find $\mathrm{E}(cX)$, we can simply apply this definition with $g(X) = cX$. Hence, using the linearity of the series,

$$\mathrm{E}(cX) = \sum_{x=0}^{\infty} cx\mathrm{pmf}(x) = c\sum_{x=0}^{\infty} x\mathrm{pmf}(x).$$

We know from (10) that

$$\sum_{x=0}^{\infty} x\mathrm{pmf}(x) = \mathrm{E}(X),$$

with $g(X) = X$. Therefore,

$$\mathrm{E}(cX) = c\sum_{x=0}^{\infty} x\mathrm{pmf}(x) = c\mathrm{E}(X),$$

as desired.

We can perform a similar calculation to find $\mathrm{Var}(X)$. Using the fact that $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$ and the previous result, we have that

$$\begin{aligned}\mathrm{Var}(cX) &= \mathrm{E}((cX)^2) - (\mathrm{E}(cX))^2 \\ &= \mathrm{E}(c^2X^2) - c^2\mathrm{E}(X)^2. \end{aligned} \tag{11}$$

To find $\mathrm{E}(c^2X^2)$, we apply (10) with $g(X) = c^2X^2$ so that, using the linearity of the series,

$$\mathrm{E}(c^2X^2) = \sum_{x=0}^{\infty} c^2x^2\mathrm{pmf}(x) = c^2\sum_{x=0}^{\infty} x^2\mathrm{pmf}(x).$$

Note that

$$\sum_{x=0}^{\infty} x^2\mathrm{pmf}(x) = \mathrm{E}(X^2)$$

using (10) with $g(X) = X^2$.

Thus,

$$\mathrm{E}(c^2X^2) = c^2 \sum_{x=0}^{\infty} x^2 \mathrm{pmf}(x) = c^2\mathrm{E}(X^2). \tag{12}$$

Therefore, combining (11) and (12), we have

$$\begin{aligned}\mathrm{Var}(cX) &= \mathrm{E}((cX)^2) - (\mathrm{E}(cX))^2 \\ &= c^2\mathrm{E}(X^2) - c^2\mathrm{E}(X)^2 \\ &= c^2(\mathrm{E}(X^2) - \mathrm{E}(X)^2) = c^2\mathrm{Var}(X),\end{aligned}$$

as desired.

Now suppose that X is a continous random variable. Then X has probability density function $\mathrm{pdf}(x)$ and

$$\mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\mathrm{pdf}(x)\,\mathrm{d}x. \tag{13}$$

Thus, to find $\mathrm{E}(cX)$, we can simply apply this definition with $g(X) = cX$. Hence, using the linearity of the integral,

$$\mathrm{E}(cX) = \int_{-\infty}^{\infty} cx\mathrm{pdf}(x)\,\mathrm{d}x = c\int_{-\infty}^{\infty} x\mathrm{pdf}(x)\,\mathrm{d}x.$$

We know from (13) that

$$\int_{-\infty}^{\infty} x\mathrm{pdf}(x)\,\mathrm{d}x = \mathrm{E}(X),$$

with $g(X) = X$. Therefore,

$$\mathrm{E}(cX) = c\int_{-\infty}^{\infty} x\mathrm{pdf}(x)\,\mathrm{d}x = c\mathrm{E}(X),$$

as desired.

We can perform a similar calculation to find $\mathrm{Var}(X)$. Using the fact that $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$ and the previous result, we have that

$$\begin{aligned}\mathrm{Var}(cX) &= \mathrm{E}((cX)^2) - (\mathrm{E}(cX))^2 \\ &= \mathrm{E}(c^2X^2) - c^2\mathrm{E}(X)^2.\end{aligned} \tag{14}$$

To find $\mathrm{E}(c^2X^2)$, we apply (13) with $g(X) = c^2X^2$ so that, using the linearity of the integral,

$$\mathrm{E}(c^2X^2) = \int_{-\infty}^{\infty} c^2x^2\mathrm{pdf}(x)\,\mathrm{d}x = c^2\int_{-\infty}^{\infty} x^2\mathrm{pdf}(x)\,\mathrm{d}x.$$

Note that

$$\int_{-\infty}^{\infty} x^2 \mathrm{pdf}(x)\,\mathrm{d}x = \mathrm{E}(X^2)$$

using (13) with $g(X) = X^2$.

Thus,

$$\mathrm{E}(c^2X^2) = c^2 \int_{-\infty}^{\infty} x^2 \mathrm{pdf}(x)\,\mathrm{d}x = c^2\mathrm{E}(X^2). \tag{15}$$

Therefore, combining (14) and (15), we have

$$\begin{aligned} \mathrm{Var}(cX) &= \mathrm{E}((cX)^2) - (\mathrm{E}(cX))^2 \\ &= c^2\mathrm{E}(X^2) - c^2\mathrm{E}(X)^2 \\ &= c^2(\mathrm{E}(X^2) - \mathrm{E}(X)^2) = c^2\mathrm{Var}(X), \end{aligned}$$

as desired. □

Problem 4. Let X_1 and X_2 be random variables. Show that, in the discrete and continuous case for X_1 and X_2, $\mathrm{E}(X_1 + X_2) = \mathrm{E}(X_1) + \mathrm{E}(X_2)$.

Solution. Suppose first that X_1 and X_2 are discrete random variables. Then X_1 and X_2 have joint probability mass function $\mathrm{pmf}(\boldsymbol{x_j}) = \mathrm{pmf}(x_{j_1}, x_{j_2})$.

Note, if $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^\intercal$, then

$$\mathrm{E}(g(\boldsymbol{X})) = \sum_{j_1} \cdots \sum_{j_n} g(x_{j_1}, \ldots, x_{j_n})\mathrm{pmf}(x_{j_1}, \ldots, x_{j_n}) \tag{16}$$

We can use (16) with $g(\boldsymbol{X}) = X_1 + X_2$ so that, due to the linearity of the series,

$$\begin{aligned} \mathrm{E}(X_1 + X_2) &= \sum_{j_1}\sum_{j_2}(x_{j_1} + x_{j_2})\mathrm{pmf}(x_{j_1}, x_{j_2}) \\ &= \sum_{j_1}\sum_{j_2}(x_{j_1}\mathrm{pmf}(x_{j_1}, x_{j_2}) + x_{j_2}\mathrm{pmf}(x_{j_1}, x_{j_2})) \\ &= \sum_{j_1}\left(\sum_{j_2} x_{j_1}\mathrm{pmf}(x_{j_1}, x_{j_2}) + \sum_{j_2} x_{j_2}\mathrm{pmf}(x_{j_1}, x_{j_2})\right) \\ &= \sum_{j_1}\sum_{j_2} x_{j_1}\mathrm{pmf}(x_{j_1}, x_{j_2}) + \sum_{j_1}\sum_{j_2} x_{j_2}\mathrm{pmf}(x_{j_1}, x_{j_2}) \end{aligned} \tag{17}$$

Note that the left sum in (17), by virtue of (16), is $\mathrm{E}(X_1)$ with $g(\boldsymbol{X}) = X_1$ and similarly the right sum in (17), by virtue of (16), is $\mathrm{E}(X_2)$ with $g(\boldsymbol{X}) = X_2$. Therefore,

$$\begin{aligned}
\mathrm{E}(X_1 + X_2) &= \sum_{j_1}\sum_{j_2} x_{j_1}\mathrm{pmf}(x_{j_1}, x_{j_2}) + \sum_{j_1}\sum_{j_2} x_{j_2}\mathrm{pmf}(x_{j_1}, x_{j_2}) \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2),
\end{aligned}$$

as desired.

Now suppose that X_1 and X_2 are continuous random variables. Then X_1 and X_2 have joint probability density function $\mathrm{pdf}(\boldsymbol{x}) = \mathrm{pdf}(x_1, x_1)$.

Note, if $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^\intercal$, then

$$\mathrm{E}(g(\boldsymbol{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n)\mathrm{pmf}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{18}$$

We can use (18) with $g(\boldsymbol{X}) = X_1 + X_2$ so that, due to the linearity of the integral,

$$\begin{aligned}
\mathrm{E}(X_1 + X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 + x_2)\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1\mathrm{pdf}(x_1, x_2) + x_2\mathrm{pdf}(x_1, x_2))\, \mathrm{d}x_1\, \mathrm{d}x_2 \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x_1\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1 + \int_{-\infty}^{\infty} x_2\mathrm{pdf}(x_1, x_2))\, \mathrm{d}x_1\right) \mathrm{d}x_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \qquad (19)
\end{aligned}$$

Note that the left integral in (19), by virtue of (18), is $\mathrm{E}(X_1)$ with $g(\boldsymbol{X}) = X_1$ and similarly, the right sum in (19), by virtue of (18), is $\mathrm{E}(X_2)$ with $g(\boldsymbol{X}) = X_2$. Therefore,

$$\begin{aligned}
\mathrm{E}(X_1 + X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2\mathrm{pdf}(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \\
&= \mathrm{E}(X_1) + \mathrm{E}(X_2),
\end{aligned}$$

as desired.

$\square$

Problem 5. Show that $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$.

Solution. Note that by definition, $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$, where

$$\mathrm{Cov}(X, Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))).$$

Combining these facts and the linearity of the expectation operator, it is straightforward to see that

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

□

Problem 6. Let $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu} = (1, 5)^\intercal$ and $\Sigma = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$. Find $\Sigma^{-1/2}$ such that $\boldsymbol{Z} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$ has a standard normal distribution.

Solution. If we diagonalize Σ such that $\Sigma = P\Lambda P^\intercal$ then $\Sigma^{-1/2} = P\Lambda^{-1/2}P^\intercal$ is a matrix satisfying $\boldsymbol{Z} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$ such that $\boldsymbol{Z}$ has a standard normal distribution. In order to diagonalize Σ, we must first find its eigenvalues and eigenvectors. Then the matrix P formed by the unit eigenvectors of Σ and Λ formed as the diagonal matrix consisting of the eigenvalues of Σ are the matrices needed to form the diagonalization.

So, the eigenvalues of Σ are the solutions to the characteristic equation of Σ, i.e. the solutions to

$$\begin{aligned}
c(\lambda) = \det(\lambda I - \Sigma) &= \begin{vmatrix} \lambda - 9 & 2 \\ 2 & \lambda - 6 \end{vmatrix} \\
&= (\lambda - 9)(\lambda - 6) - 4 = (\lambda - 10)(\lambda - 5) \qquad (20)
\end{aligned}$$

Hence, the roots of (20), i.e. the eigenvalues of Σ, are $\lambda_1 = 10$ and $\lambda_2 = 5$. The eigenvectors associated to λ_1 and λ_2 are, respectively, the vectors $\boldsymbol{v}_{\lambda_1}$ and $\boldsymbol{v}_{\lambda_2}$ satisfying the equation

$$(\lambda_i I - \Sigma)\boldsymbol{v}_{\lambda_i} = \mathbf{0} \quad \text{for } i = 1, 2.$$

Thus, for $\lambda_1 = 10$, we have for $\boldsymbol{v}_{\lambda_1} = (x_1, x_2)^\intercal$,

$$(\lambda_1 I - \Sigma)\boldsymbol{v}_{\lambda_1} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

By inspection, we can see that $x_1 = -2x_2$ is the solution to the above system of equations. Hence, $\boldsymbol{v}_{\lambda_1} = (-2, 1)^\intercal$ is the eigenvector associated to $\lambda_1 = 10$.

Similarly, for $\lambda_2 = 5$, we have for $\boldsymbol{v}_{\boldsymbol{\lambda_2}} = (x_1, x_2)^\intercal$,

$$(\lambda_2 I - \Sigma)\boldsymbol{v}_{\lambda_2} = \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Again, by inspection, we can see that $2x_1 = x_2$ is the solution to the above system of equations. Thus, $\boldsymbol{v}_{\lambda_2} = (1, 2)^\intercal$ is the eigenvector associated to $\lambda_2 = 5$.

We can make the eigenvectors $\boldsymbol{v}_{\lambda_1}$ and $\boldsymbol{v}_{\lambda_2}$ unit by dividing each vector by its length. Hence, we have $\boldsymbol{v}'_{\lambda_1} = (-2/\sqrt{5}, 1/\sqrt{5})^\intercal$ and $\boldsymbol{v}'_{\lambda_2} = (1/\sqrt{5}, 2/\sqrt{5})^\intercal$ as the unit eigenvectors associated to $\boldsymbol{v}_{\lambda_1}$ and $\boldsymbol{v}_{\lambda_2}$. Therefore, the matrices $P = \begin{pmatrix} \boldsymbol{v}'_{\lambda_1} & \boldsymbol{v}'_{\lambda_2} \end{pmatrix}$ and $\Lambda = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}$ form the diagonalization of $\Sigma = P\Lambda P^\intercal$.

Thus, $\Sigma^{-1/2} = P\Lambda^{-1/2}P^\intercal$ is the matrix satisfying $\boldsymbol{Z} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$ where $\boldsymbol{Z}$ has a standard normal distribution.

Therefore,

$$
\begin{aligned}
\Sigma^{-1/2} &= P\Lambda^{-1/2}P^\intercal \\
&= \begin{pmatrix} \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{10}} & 0 \\ 0 & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{-2}{\sqrt{50}} & \frac{1}{5} \\ \frac{1}{\sqrt{50}} & \frac{2}{5} \end{pmatrix} \begin{pmatrix} \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{4}{5\sqrt{10}} + \frac{1}{5\sqrt{5}} & \frac{-2}{5\sqrt{10}} + \frac{2}{5\sqrt{5}} \\ \frac{-2}{5\sqrt{10}} + \frac{2}{5\sqrt{5}} & \frac{1}{5\sqrt{10}} + \frac{4}{5\sqrt{5}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{4+\sqrt{2}}{5\sqrt{10}} & \frac{-2+2\sqrt{2}}{5\sqrt{10}} \\ \frac{-2+2\sqrt{2}}{5\sqrt{10}} & \frac{1+4\sqrt{2}}{5\sqrt{10}} \end{pmatrix}
\end{aligned}
$$

is the desired matrix.

$\square$

Problem 7. Let $\boldsymbol{X} = (X_1, X_2) \sim N(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = (\mu_1, \mu_2)^\intercal$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Show that the probability density function of $\boldsymbol{X}$, $f(\boldsymbol{x})$, for $\boldsymbol{x} = (x_1, x_2)^\intercal$ is

$$f(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{-2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

Solution. By definition, if $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$, then

$$f(\boldsymbol{x}) = (2\pi)^{-\frac{n}{2}}(\det\Sigma)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \tag{21}$$

Note, since $n = 2$ in this case, we need only find $\det\Sigma$, Σ^{-1}, and $\boldsymbol{x} - \boldsymbol{\mu}$ to find the desired probability density function $f(\boldsymbol{x})$.

Since $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ is a 2×2 matrix, it's clear that $\det \Sigma = \sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2$. Again, since Σ is a 2×2 matrix, we can use $\det \Sigma$ to calculate Σ^{-1}. Thus,

$$\Sigma^{-1} = \begin{pmatrix} \dfrac{\sigma_2^2}{\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2} & -\dfrac{\rho\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2} \\ -\dfrac{\rho\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2} & \dfrac{\sigma_1^2}{\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2} \end{pmatrix}.$$

Finally, $(\boldsymbol{x} - \boldsymbol{\mu}) = (x_1 - \mu_1, x_2 - \mu_2)^\intercal$. Using these computations and the definition in (21), it is straightforward that

$$\begin{aligned} f(\boldsymbol{x}) &= (2\pi)^{-\frac{n}{2}}(\det \Sigma)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\intercal \Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right\} \\ &= (2\pi)^{-1}(\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x_1 - \mu_1, x_2 - \mu_2)^\intercal \Sigma^{-1} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}\right\} \\ &= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1 - \rho^2)}} \exp\left\{-\frac{(\sigma_2^2(x_1 - \mu_1)^2 - 2\rho\sigma_1\sigma_2(x_1 - \mu_1)(x_2 - \mu_2) + \sigma_1^2(x_2 - \mu_2)^2)}{2(\sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2)}\right\} \\ &= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1 - \rho^2)}} \exp\left\{-\frac{1}{2(1 - p^2)}\left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{-2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}\right]\right\}, \end{aligned}$$

as desired.

$\square$