Question 1

(a) Prove that for any two random variables X and Y,

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

(b) Prove that for random variables $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ that

$$\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathrm{Cov}\,(X_i, Y_j)$$

Solution to Question 1

Part (a):

By definition, if we write $\mathrm{E}\,(X) = \mu_X$ and $\mathrm{E}\,(Y) = \mu_Y$, then using the linearity of expectation,

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

Part (b):

We use Part (a).

$$\begin{aligned}
\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) &= \mathrm{E}\left[\left(\sum_{i=1}^{n} a_i X_i\right)\left(\sum_{j=1}^{m} b_j Y_j\right)\right] - \mathrm{E}\left[\sum_{i=1}^{n} a_i X_i\right]\mathrm{E}\left[\sum_{j=1}^{m} b_j Y_j\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j X_i Y_j\right] - \mathrm{E}\left[\sum_{i=1}^{n} a_i X_i\right]\mathrm{E}\left[\sum_{j=1}^{m} b_j Y_j\right]
\end{aligned}$$

and using the linearity of expectation,

$$\begin{aligned}
\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathrm{E}\,[X_i Y_j] - \left(\sum_{i=1}^{n} a_i \mathrm{E}\,[X_i]\right)\left(\sum_{j=1}^{m} b_j \mathrm{E}\,[Y_j]\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathrm{E}\,[X_i Y_j] - \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathrm{E}\,[X_i]\,\mathrm{E}\,[Y_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \left(\mathrm{E}\,[X_i Y_j] - \mathrm{E}\,[X_i]\,\mathrm{E}\,[Y_j]\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \mathrm{Cov}\,(X_i, Y_j,)
\end{aligned}$$

Question 2

Young's inequality states that if a and b are non-negative real numbers, and p and q are any positive numbers such that

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab.$$

Use Young's inequality to prove Hölder's Inequality: Let X and Y be random variables and let p and q be two positive numbers satisfying $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$|\mathrm{E}\left(XY\right)| \leq \left(\mathrm{E}\left(|X|^p\right)\right)^{1/p} \left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}.$$

(a) If Z is a non-negative random variable, i.e. $Z \geq 0$, prove that $\mathrm{E}\left(Z\right) \geq 0$.

(b) Prove that $\mathrm{E}\left(|XY|\right) \geq 0$.

(c) Prove that $|\mathrm{E}\left(XY\right)| \leq \mathrm{E}\left(|XY|\right)$.

(d) Use Young's inequality to prove $\mathrm{E}\left(|XY|\right) \leq \left(\mathrm{E}\left(|X|^p\right)\right)^{1/p} \left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}$.

(e) Conclude that Hölder's Inequality is true.

(f) Use Hölder's Inequality to prove the Cauchy-Schwarz Inequality: $|\mathrm{E}\left(XY\right)| \leq (\mathrm{E}(|X|^2))^{1/2}(\mathrm{E}(|Y|^2))^{1/2}$.

(g) Use the Cauchy-Schwarz inequality to prove Theorem 4.1.6: $|\mathrm{Cov}\left(X, Y\right)| \leq \sigma_X \sigma_Y$, where σ_X^2 and σ_Y^2 are the variances of X and Y, respectively.

Solution to Question 2

Part (a):

We consider the case that Z is a continuous random variable; the discrete case is similar.

Let f be the probablity density function of Z. Then $f(z) \geq 0$ for any $z \in \mathbb{R}$. Since $Z \geq 0$, the support of f is a subset of $\{z \in \mathbb{R} : z \geq 0\}$. Now,

$$\begin{aligned}
&\forall z \geq 0, f(z) \geq 0 \\
\Rightarrow &\forall z \geq 0, z f(z) \geq 0 \\
\Rightarrow &\int_0^\infty z f(z) \mathrm{d}z \geq 0 \\
\Rightarrow &\mathrm{E}\left(Z\right) \geq 0
\end{aligned}$$

Part (b):

For any random variables X and Y, define the random variable $Z = |XY|$. Then Z is a non-negative random variable, and from Part (a) $\mathrm{E}\left(|XY|\right) = \mathrm{E}\left(Z\right) \geq 0$.

Part (c):

First, we note that $-|XY| \leq XY \leq |XY|$. Now,

$$\begin{aligned}
&-|XY| \leq XY \leq |XY| \\
\Rightarrow &-\mathrm{E}\left(|XY|\right) \leq \mathrm{E}\left(XY\right) \leq \mathrm{E}\left(|XY|\right)
\end{aligned}$$

The second line follows from (for example) considering $f_{X,Y}$ to be the joint probability density function of X and Y and

$$\begin{aligned}
&\forall x, y \in \mathbb{R}, \qquad xy \leq |xy| \\
\Rightarrow &\forall x, y \in \mathbb{R}, \qquad xy f_{X,Y}(x,y) \leq |xy| \, f_{X,Y}(x,y) \\
\Rightarrow &\int_{-\infty}^{\infty} xy f_{X,Y}(x,y) \mathrm{d}x \mathrm{d}y \leq \int_{-\infty}^{\infty} |xy| \, f_{X,Y}(x,y) \mathrm{d}x \mathrm{d}y \\
\Rightarrow &\mathrm{E}\left(XY\right) \leq \mathrm{E}\left(|XY|\right)
\end{aligned}$$

This proves the second inequality; the first inequality $-\mathrm{E}\left(|XY|\right) \leq \mathrm{E}\left(XY\right)$ follows similarly, and using $\mathrm{E}\left(-|XY|\right) = -\mathrm{E}\left(|XY|\right)$.

Next, since for any $a \in \mathbb{R}$ and any $b \in \mathbb{R}, b \geq 0$,

$$\begin{aligned}
&-b \leq a \leq b \\
\Rightarrow &|a| \leq b
\end{aligned}$$

and since from Part (b) we have $\mathrm{E}\left(|XY|\right) \geq 0$, then

$$\begin{aligned}
&-\mathrm{E}\left(|XY|\right) \leq \mathrm{E}\left(XY\right) \leq \mathrm{E}\left(|XY|\right) \\
\Rightarrow &\left|\mathrm{E}\left(XY\right)\right| \leq \mathrm{E}\left(|XY|\right)
\end{aligned}$$

Part (d):

Starting with Young's Inequality,

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab.$$

for p, q such that $\frac{1}{p} + \frac{1}{q} = 1$. Now define

$$a = \frac{|X|}{\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}}, \qquad b = \frac{|Y|}{\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}}.$$

Then, plugging these values into Young's Inequality,

$$\frac{1}{p}\frac{|X|^p}{\mathrm{E}\left(|X|^p\right)} + \frac{1}{q}\frac{|Y|^q}{\mathrm{E}\left(|Y|^q\right)} \geq \frac{|XY|}{\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}}$$

Now, taking expectations of both sides and using the linearity of expectation,

$$\begin{aligned}
&\frac{1}{p}\frac{\mathrm{E}\left(|X|^p\right)}{\mathrm{E}\left(|X|^p\right)} + \frac{1}{q}\frac{\mathrm{E}\left(|Y|^q\right)}{\mathrm{E}\left(|Y|^q\right)} \geq \frac{\mathrm{E}\left(|XY|\right)}{\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}} \\
\Rightarrow &\frac{1}{p}\left(1\right) + \frac{1}{q}\left(1\right) \geq \frac{\mathrm{E}\left(|XY|\right)}{\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}} \\
\Rightarrow &1 \geq \frac{\mathrm{E}\left(|XY|\right)}{\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q}} \\
\Rightarrow &\left(\mathrm{E}\left(|X|^p\right)\right)^{1/p}\left(\mathrm{E}\left(|Y|^q\right)\right)^{1/q} \geq \mathrm{E}\left(|XY|\right)
\end{aligned}$$

as required.

Part (e):

Putting together Parts (c) and (d),

$$|\mathrm{E}\,(XY)| \leq \mathrm{E}\,(|XY|) \leq (\mathrm{E}\,(|X|^p))^{1/p}(\mathrm{E}\,(|Y|^q))^{1/q}$$

which proves Hölder's Inequality.

Part (f):

Taking $p = q = 2$, Hölder's Inequality becomes

$$|\mathrm{E}\,(XY)| \leq (\mathrm{E}(|X|^2))^{1/2}(\mathrm{E}(|Y|^2))^{1/2}$$

Part (g):

If we define μ_X and μ_Y to be the means of X and Y, respectively, and σ_X^2 and σ_Y^2 to be the variances of X and Y, respectively, then starting with the Cauchy-Schwarz inequality for the random variables $X - \mu_X$ and $Y - \mu_Y$,

$$|\mathrm{Cov}\,(X,Y)| = |\mathrm{E}[(X-\mu_X)\,(Y-\mu_Y)]| \leq (\mathrm{E}(|X-\mu_X|^2))^{1/2}(\mathrm{E}(|Y-\mu_Y|^2))^{1/2} = \sigma_X\sigma_Y.$$

Alternative proof:

An alternative approach is to prove the Cauchy-Schwarz Inequality in terms of inner products, i.e.

$$|\langle X, Y\rangle|^2 \leq \langle X, X\rangle\langle Y, Y\rangle.$$

and then define the inner product $\langle X, Y\rangle = \mathrm{E}\,(XY)$ (you will need to show this is an inner product), and then defining μ_X and μ_Y as the means of X and Y, respectively, and σ_X^2 and σ_Y^2 as the variances of X and Y, respectively,

$$\begin{aligned}
(\mathrm{Cov}\,(X,Y))^2 &= |\mathrm{Cov}\,(X,Y)|^2 \\
&= |\mathrm{E}[(X-\mu_X)\,(Y-\mu_Y)]|^2 \\
&= |\langle X-\mu_X, Y-\mu_Y\rangle|^2 \\
&\leq \langle X-\mu_X, X-\mu_X\rangle\langle Y-\mu_Y, Y-\mu_Y\rangle \qquad \text{(Cauchy-Schwarz Inequality)} \\
&= \mathrm{E}[(X-\mu_X)^2]\mathrm{E}[(Y-\mu_Y)^2] \\
&= \mathrm{Var}\,(X)\,\mathrm{Var}\,(Y) \\
&= \sigma_X^2\sigma_Y^2 \\
\Rightarrow \mathrm{Cov}\,(X,Y) &\leq \sigma_X\sigma_Y
\end{aligned}$$