Introductory Econometrics

Tutorial 3 (suggested answers)

<u>PART A:</u> To be done before you attend the tutorial. The solutions will be made available at the end of the week.

This homework is a review of statistical concepts of mean, variance, covariance and correlation.

1. No. $E(X) = \sum_{i=1}^{\#\ of\ possible\ outcomes} x_i p_i$. We are given all possible outcomes, but not the probability of each outcome. We need the probabilities of each outcome to be able to determine $E(X)$.

2.

$$
\begin{aligned}
E(2X - Y) &= 2E(X) - E(Y) = 2 \times 3 - 4 = 2. \\
Var(2X - Y) &= 2^2 Var(X) + (-1)^2 Var(Y) + 2 \times 2 \times (-1) \times Cov(X, Y) \\
&= 4Var(X) + Var(Y) = 4 \times 4 + 9 = 25 \ \ [\text{since } Cov(X, Y) = 0].
\end{aligned}
$$

3. From the joint pdf, we can deduce the marginal pdf of the cash rate by summing the probabilities in each row, and we can deduce the marginal pdf of the inflation rate by summing the probabilities in each column. Then using these, we can compute the expected value of the cash rate, and the expected value of the inflation rate.

cash rate ↓	Pr
1.25%	0.25
1.50%	0.50
1.75%	0.25

$$E(\text{cah rate}) = 1.25 \times 0.25 + 1.50 \times 0.50 + 1.75 \times 0.25 = 1.50\%$$

inflation rate ⟶	1%	2%	3%
Pr	0.40	0.50	0.10

$$E(\text{inflation rate}) = 1 \times 0.40 + 2 \times 0.50 + 3 \times 0.10 = 1.7\%$$

To compute the conditional mean of inflation rate conditional of cash rate being 1.50%, we need the conditional distribution of inflation rate when cash rate is 1.50%.

inflation rate ⟶	1%	2%	3%
Pr(. | cash rate = 1.50)	0.15/0.50 = 0.30	0.30/0.50 = 0.60	0.05/0.50 = 0.10

$$E(\text{inflation rate} \mid \text{cash rate } = 1.50) = 1 \times 0.30 + 2 \times 0.60 + 3 \times 0.10 = 1.80\%$$

Do not forget to bring your answers to PART A and a copy of the tutorial questions to your tutorial.

PART B: You do not need to hand this part in. It will be covered in the tutorial. It is still a good idea to attempt these questions before the tutorial.

1.

$$
\begin{aligned}
Cov\left(X,Y\right) &= E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right] = E\left[\left(X-\mu_X\right)Y-\left(X-\mu_X\right)\mu_Y\right] \\
&= E\left[\left(X-\mu_X\right)Y\right]-E[\left(X-\mu_X\right)\mu_Y] \\
&= E\left[\left(X-\mu_X\right)Y\right]-\mu_Y\underbrace{E\left(X-\mu_X\right)}_{0}=E\left[\left(X-\mu_X\right)Y\right] \\
&= E\left(XY-\mu_X Y\right)=E\left(XY\right)-\mu_X E\left(Y\right)=E\left(XY\right)-\mu_X\mu_Y.
\end{aligned}
$$

The last expression $E\left(XY\right)-\mu_X\mu_Y$ cannot be simplified further to $E\left(X\right)E\left(Y\right)-\mu_X\mu_Y$ because $E\left(XY\right)\neq E\left(X\right)E\left(Y\right)$ in general. (Those are equal when X and Y are independent.)

2. *Let $E\left(X\right)=\mu$ and $Var\left(X\right)=\sigma^2$. Then when we bet all of one dollar at once, the expected winning is $E\left(4X\right)=4\mu$, and variance of this bet is $Var\left(4X\right)=4^2Var\left(X\right)=16\sigma^2$. If we use our dollar to bet X four different times, we have the expected winning of $E\left(X_1+X_2+X_3+X_4\right)=\mu+\mu+\mu+\mu=4\mu$, and the variance of the outcome of this strategy is $Var\left(X_1+X_2+X_3+X_4\right)=Var\left(X_1\right)+Var\left(X_2\right)+Var\left(X_3\right)+Var\left(X_4\right)=\sigma^2+\sigma^2+\sigma^2+\sigma^2=4\sigma^2$. So, Farshid's Mum is right because the first strategy has the same expected return as the second one, but has a much larger risk.*

3. *Let's denote the unknown mean and variance of the random variable μ and σ^2. Our sample of one observation is X. Its mean is obviously $E\left(X\right)=\mu$ and its variance is $Var\left(X\right)=\sigma^2$. A random sample of 4 observations canbe denoted by $\{X_1,X_2,X_3,X_4\}$ and its average is $\overline{X}=\frac{1}{4}\left(X_1+X_2+X_3+X_4\right)=\frac{1}{4}\sum_{i=1}^{i=4}X_i$. The expected value of $\overline{X}$ is*

$$
\begin{aligned}
E\left(\overline{X}\right) &= E\left[\frac{1}{4}\left(X_1+X_2+X_3+X_4\right)\right]=\frac{1}{4}E\left(X_1+X_2+X_3+X_4\right) \\
&= \frac{1}{4}\left[E\left(X_1\right)+E\left(X_2\right)+E\left(X_3\right)+E\left(X_4\right)\right] \\
&= \frac{1}{4}\left[\mu+\mu+\mu+\mu\right]=\mu
\end{aligned}
$$

and its variance is

$$
\begin{aligned}
Var\left(\overline{X}\right) &= Var\left[\frac{1}{4}\left(X_1+X_2+X_3+X_4\right)\right]=\frac{1}{16}Var\left(X_1+X_2+X_3+X_4\right) \\
&= \frac{1}{16}\left[Var\left(X_1\right)+Var\left(X_2\right)+Var\left(X_3\right)+Var\left(X_4\right)\right] \qquad \text{because the sample is random} \\
&= \frac{1}{16}\left[\sigma^2+\sigma^2+\sigma^2+\sigma^2\right]=\frac{1}{16}\times 4\sigma^2=\frac{\sigma^2}{4}
\end{aligned}
$$

So both a sample of one observation and the average of the sample of 4 observations have the same expected value (which is the parameter that we want to estimate), but the average of 4 observations is less risky (i.e. it is less likely to be very far from μ). This is very similar to Farshid's Mum's Theorem. There we were comparing $4X$ and $\left(X_1+X_2+X_3+X_4\right)$, here we are comparing X and $\frac{1}{4}\left(X_1+X_2+X_3+X_4\right)$. Same principle: it is safer to diversify and not depend only on a single draw from the distribution.

4.

$$X = \mathbf{p}'\mathbf{z}$$

$$\begin{aligned} E(X) &= \mathbf{p}'E(\mathbf{z}) = \begin{pmatrix} 0.2 & 0.3 & 0.5 \end{pmatrix} \begin{pmatrix} 1.0 \\ 0.6 \\ 0.8 \end{pmatrix} \\ &= 0.2 \times 1.0 + 0.3 \times 0.6 + 0.5 \times 0.8 = 0.78 \end{aligned}$$

$$\begin{aligned} Var(X) &= Var(\mathbf{p}'\mathbf{z}) = \mathbf{p}'Var(\mathbf{z})\,\mathbf{p} \\ &= \begin{pmatrix} 0.2 & 0.3 & 0.5 \end{pmatrix} \begin{bmatrix} 94 & 1 & 8 \\ 1 & 20 & 5 \\ 8 & 5 & 21 \end{bmatrix} \begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix} \\ &= \begin{pmatrix} 23.1 & 8.7 & 13.6 \end{pmatrix} \begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix} = 14.03 \end{aligned}$$

Note that the expected return of this portfolio (0.78) *is almost the same as the rate of return to Wesfarmers shares* (0.8)*, but its variance is well below the variance of return to any of the three assets (*94 *for Qantas,* 20 *for Telstra and* 21 *for Wesfarmers shares. Even with a portfolio comprising 100 assets, a software package that can handle matrices can compute these very quickly and efficiently. Of course the means, variances and covariances of asset returns need to be estimated from historical data first.*