Practical 5 Solutions

Question 1

(a)

$$\begin{aligned} f(x) &= \int_0^1 f(x,y)dy \\ &= \int_0^1 12xy(1-x)dy \\ &= \int_0^1 12xy - 12x^2y dy \\ &= \left[\frac{12xy^2}{2} - \frac{12x^2y^2}{2}\right]_0^1 \\ &= \frac{12x}{2} - \frac{12x^2}{2} \\ &= 6x - 6x^2 = 6x(1-x). \end{aligned}$$

(b) To determine if X and Y are independent we see if we can factorise the joint density as

$$f(x,y) = f(x)f(y)$$

We have

$$f(x,y) = 12xy(1-x) = 2y(6x(1-x)) = f(y)f(x),$$

so X and Y are indeed independent.

(c)

$$\begin{aligned} E(X) &= \int_0^1 x \times 6x(1-x)dx \\ &= 6\int_0^1 x^2 - x^3 dx \\ &= 6\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 \\ &= 6\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{1}{2} \end{aligned}$$

(d)

$$\begin{aligned} E(X^2) &= \int_0^1 x^2 \times 6x(1-x)dx \\ &= 6\int_0^1 x^3 - x^4 dx \\ &= 6\left[\frac{x^4}{4} - \frac{x^5}{5}\right]_0^1 \\ &= 6\left(\frac{1}{4} - \frac{1}{5}\right) = \frac{3}{10}. \end{aligned}$$

Therefore $Var(X)$ can be found as

$$Var(X) = E(X^2) - E(X)^2 = \frac{3}{10} - \left(\frac{1}{2}\right)^2 = \frac{1}{20}.$$

Question 2

(a) We have $X \sim Bin(100, 0.5)$ and require $P(X > 60)$. So we first calculate the mean and variance of X. $E(X) = n\pi = 100(0.5) = 50$ and $Var(X) = n\pi(1-\pi) = 100 \times 0.5 \times 0.5 = 25$. So we approximate the probability using the normal distribution, where we use Z to indicate a standard Normal random variable,

$$\begin{aligned}
P(X > 60) = 1 - P(X \leq 60) &= 1 - P\left(Z \leq \frac{60.5 - 50}{\sqrt{25}}\right) \\
&= 1 - P(Z \leq 2.1) \\
&= 1 - \Phi(2.1) \\
&= 1 - 0.982 = 0.018
\end{aligned}$$

(b) We have $X \sim Bin(48, 0.75)$ and require $P(30 < X < 39)$. Again we first calculate the mean and variance of X. $E[X] = n\pi = 48 \times 0.75 = 36$ and $Var(x) = n\pi(1-\pi) = 48 \times 0.75 \times 0.25 = 9$. Thus

$$\begin{aligned}
P(30 < X < 39) &= P(31 \leq X \leq 38) \\
&= P(30.5 \leq Y \leq 38.5) \text{ where } Y \text{ is our normal approximation} \\
&= P(Y < 38.5) - P(Y < 30.5) \\
&= P\left(Z \leq \frac{38.5 - 36}{3}\right) - P\left(Z \leq \frac{30.5 - 36}{3}\right) \\
&= P(Z \leq 0.833) - P(Z \leq -1.833) \\
&= \Phi(0.833) - \Phi(-1.833) \\
&= 0.798 - 0.033 = 0.764
\end{aligned}$$

(c) In this case $X \sim Poisson(30)$ and we require $P(X \leq 20)$. Again we first calculate the mean and variance of X.

$E(X) = \mu = 30$ and $Var(X) = \mu = 30$. Approximating by a normal we have

$$\begin{aligned}
Pr(X \leq 20) &= P\left(Z \leq \frac{20.5 - 30}{30^{0.5}}\right) \\
&= P(Z \leq -1.734) \\
&= \Phi(-1.734) \\
&= 0.0414
\end{aligned}$$

(d) This is a slightly more complicated example where we need to find $P(X_1 + X_2 \leq 40)$ where X_1, X_2 are iid $Poisson(30)$.

We use the formula from the lecture to find the mean of $X_1 + X_2$ as follows:

$$E(X_1 + X_2) = E(X_1) + E(X_2) = 30 + 30 = 60$$

Since we are given that X_1 and X_2 are independent we can use the formula for variance to find the variances of the sum:

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) = 30 + 30 = 60$$

The sum of two independent Poisson random variables is also a Poisson variable, although we have not proved this. Having found these values we can use the usual method to calculate the probabilities:

$$
\begin{aligned}
P(X_1 + X_2 \leq 40) &= P\left(Z \leq \frac{40.5 - 60}{60^{0.5}}\right) \\
&= P(Z \leq -2.517) \\
&= \Phi(-2.517) \\
&= 0.006
\end{aligned}
$$

Question 3

(a) The marginal distribution is $N(85, 22.2)$.

(b) The correlation between leg length at age 6 and age 2 is

$$\rho = \frac{13.7}{\sqrt{22.2} \times \sqrt{29}} = 0.54.$$

The conditional distribution of leg length at age 6 conditional on leg length at age 2 is normal, with conditional expectation and variance:

$$
\begin{aligned}
E(length_6|length_2) &= 114 + \frac{0.54 \times \sqrt{29.0}}{\sqrt{22.2}}(length_2 - 85) \\
Var(length_6|length_2) &= 29.0(1 - 0.54^2) = 20.5.
\end{aligned}
$$

Additional: Question 4

To show that U and V are uncorrelated we need to show that they have zero covariance:

$$
\begin{aligned}
Cov(U, V) = Cov(X - Y, X + Y) &= Cov(X, X) + Cov(X, Y) - Cov(Y, X) - Cov(Y, Y) \\
&= Var(X) - Var(Y),
\end{aligned}
$$

since $Cov(X, Y) = Cov(Y, X)$, and this is equal to zero if X and Y have equal variance.

Additional: Question 5

First we note that since $Var(X) = 1$, $E(X^2) = 1$. Then the covariance is equal to

$$
\begin{aligned}
Cov(X, X^2) &= E(X(X^2 - E(X^2))) \\
&= E(X(X^2 - 1)) \\
&= E(X^3 - X) \\
&= 0 - 0 = 0,
\end{aligned}
$$

using the fact we are given that $E(X^3) = 0$ when $X \sim N(0, 1)$.

Additional: Question 6

This question can be tackled fairly simply by using the bivariate normal joint density function. Below, we show how to tackle the question directly from the general MVN density.

We first recall the joint density function of the MVN:

$$f(\mathbf{x}) = |2\pi\Sigma|^{-1/2}\exp(-(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})/2)$$

Our aim is to show that when $\sigma_{12} = 0$, this factorises into the product of the marginal densities $f(x_1) \times f(x_2)$.

We first consider the determinant of the matrix $2\pi\Sigma$:

$$\left|\begin{pmatrix} 2\pi\sigma_1^2 & 0 \\ 0 & 2\pi\sigma_2^2 \end{pmatrix}\right| = 4\pi^2\sigma_1^2\sigma_2^2$$

When this is raised to the power $-1/2$ this is equal to the product of the corresponding parts in the univariate normal distributions. Thus all that remains is to show that exponential part factorises. To do this we first find Σ^{-1}:

$$\begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2}\begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}.$$

Then we evaluate the argument to the exponential (ignoring the $-1/2$ factor):

$$\begin{aligned}
(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) &= \frac{1}{\sigma_1^2\sigma_2^2}(x_1-\mu_1, x_2-\mu_2)\begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\
&= \frac{1}{\sigma_1^2\sigma_2^2}(\sigma_2^2(x_1-\mu_1), \sigma_1^2(x_2-\mu_2))\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\
&= \frac{1}{\sigma_1^2\sigma_2^2}(\sigma_2^2(x_1-\mu_1)^2 + \sigma_1^2(x_2-\mu_2)^2) \\
&= \frac{1}{\sigma_1^2}(x_1-\mu_1)^2 + \frac{1}{\sigma_2^2}(x_2-\mu_2)^2,
\end{aligned}$$

and so this factorises into the product of the exponential components from the respective marginal densities. Since the joint density factorises into this product, the X_1 and X_2 are independent when $\sigma_{12} = 0$.