1 Question 1

1.1 Part a

$(X_t)_t$ is an ergodic Markov chain as P^k for $k = 4$ results in a matrix with strictly positive entries. It should also be noted that the Markov chain is irreducible and state 2 is aperiodic. This confirms that $(X_t)_t$ is an ergodic Markov chain.

1.2 Part b

Using $P(X_t = j|X_{(t-k)} = i) = (P^k)_{i,j}$ from page 58 of the notes. Where $(P^k)_{i,j}$ is the entry in the ith row and jth column of the transition matrix raised to the power k. This means $P(X_5 = 4|X_2 = 1) = (P^3)_{1,4} = 0.1250$

1.3 Part c

To compute the stationary distribution π of the Markov chain, we need to solve $\pi = \pi P$ and $\pi\ (1, 1, 1, 1)^T = 1$. Solving this linear system gives:

$$\pi = (0.0909, 0.3636, 0.2424, 0.3030)$$

As the Markov chain is ergodic and π is a stationary distribution then π is unique.

1.4 Part d

A Markov chain $(X_t)_t$ with transition matrix P is said to satisfy detailed balance with distribution π if:

$$\frac{\pi_i}{\pi_j} P_{ij} = P_{ji} \quad \forall i, j \in \{1, 2, 3, 4\}$$

Through direct computation we note it does not hold for $(i, j) = (2, 3)$:

$$\frac{0.3636}{0.2424} 0.25 = 0.375 \neq 0.5$$

2 Question 2

2.1 Part a

To estimate $\mathbb{E}(X^2)$ from the given distribution $f(x) = k \exp(-2|x|^5)$ using a Metropolis-Hastings sampler, we draw a sample $(X_t)_t$ and use Monte Carlo integration to compute $\mathbb{E}(X^2)$. i.e.

$$\mathbb{E}(X^2) = \frac{1}{n} \sum_{t=1}^{n} X_t^2$$

Note: The normalising constant k does not need to be computed for the Metropolis-Hastings sampler. Using a random walk Metropolis-Hastings sampler with initial value $X_0 = 0$ and $\sigma = 0.5$, we sample from f. Then removing the first 100 observations to allow for burn-in, we compute the Monte Carlo estimate of $\mathbb{E}(X^2) = 0.2484$. To check this answer using numerical integration, we need to compute the value of the normalising constant $k = 0.6255$. We wish to use numerical integration to solve:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

Placing the integrand into the integrate function on R gives a value of 0.2458 with absolute error $< 3.4e - 06$. This means our Monte Carlo estimate using the sample from our Metropolis-Hastings sampler was within 3 decimals places of the true answer.

2.2 Part b

Constructing four samples using the parameters specified in the question, we get the following trace plot:

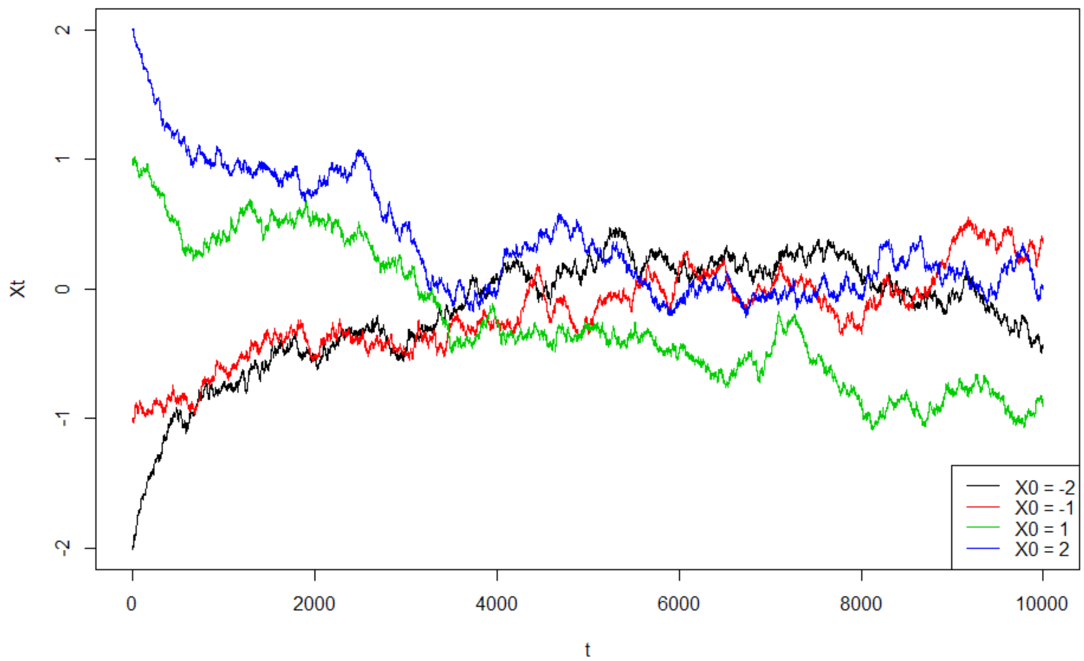

Figure 1: Trace plots of four samples with different starting values

As can be seen in Figure 1, 3 trace plots seem to overlap with each other and tend towards 0. However, one trace plot does not tend towards 0. These trace plots do not appear to converge and it is clear from the small movements shown by the trace plots that the distribution is not being fully explored. This is due to the low value of $\sigma = 0.01$. Increasing this value of σ will increase the exploration of the distribution and improve convergence.

2.3 Part c

We draw four samples using the parameters specified in the question and $X_0 = 0$. We remove the first 250 points from each sample as burn-in. The Effective Sample Size (ESS) of each sample is as follows:

Effective sample size is the size of an iid random sample from the target distribution which would be required to match the variance of our Metropolis-Hastings generated sample. This is due to the Metropolis-Hastings sampler producing a sample which inherently has correlation within and therefore degrades the quality of the sample. Due to this, we wish to choose the sample with the

σ	Effective Sample Size
0.01	5.1853
0.5	1439.1610
1	2307.5274
5	771.2703

Table 1: σ and its effect on ESS

largest Effective Sample Size, as it requires the largest iid sample to match our MCMC variance. Looking at table 1, this occurs when $\sigma = 1$ with an ESS of 2307 with n = 9750 (after burn in).

3 Question 3

Let $g(x_1, x_2)$ be the joint density of a bi-variate normal. Then we can conclude that:

$$(X_1, X_2) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.8\sqrt{2} \\ 0.8\sqrt{2} & 2 \end{bmatrix} \right)$$

To construct a Gibb's sampler for $g(x_1, x_2)$, we need to construct the conditional distribution of $X_1|X_2$ and $X_2|X_1$. For a bi-variate normal these are well known and readily available on the Multivariate normal distribution wikipedia page. The conditional distribution is given as:

$$(X_1|X_2 = x_2) \sim N\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x_2 - \mu_2) \ , \ \sigma_1^2(1 - \rho^2) \right)$$

Given the values from the question; $\mu_1, \mu_2 = 0, \sigma_1 = 1, \sigma_2 = \sqrt{2}$ and $\rho = 0.8$, we get:

$$(X_1|X_2 = x_2) \sim N\left(\frac{0.8(x_2)}{\sqrt{2}} \ , \ (1 - 0.8^2) \right) \tag{1}$$

$$(X_2|X_1 = x_1) \sim N\left(0.8\sqrt{2}x_1 \ , \ 2(1 - 0.8^2) \right) \tag{2}$$

Using results (1) and (2), we can construct a Gibb's sampler as follows:

1. Choose initial values x_1^0, x_2^0 for x_1 and x_2 respectively. For this example, I chose $(x_1^0, x_2^0) = (0, 0)$ as $g(x_1, x_2)$ is symmetric around $(0, 0)$.
2. Sample $(X_1^t | X_2 = X_2^{t-1})$ using result (1)
3. Sample $(X_2^t | X_1 = X_1^t)$ using result (2)
4. repeat steps 2 and 3 for each $t = 1, \ldots, 10000$

We construct a sample from a bi-variate normal using the Gibb's sampler above. We remove the first 50 observations to allow for burn-in. We can estimate $P((X_1, X_2) \in [0, 1] \times [0, 2])$ by counting the number of occurrences seen in the sample and dividing by 9950. This gives an estimate of 0.2270.

4 Question 4

4.1 Part a

We can construct the 95% confidence intervals for $T(X)$ using the result given on page 42 of the notes and 10000 bootstrapped samples of X, X^*:

$$[T(\hat{P}) - c_2^*\sigma(\hat{P}), T(\hat{P}) - c_1^*\sigma(\hat{P})] \tag{3}$$

where $\hat{P}$ is the original data sample,X,
$T(P)$ is a function of P,
$\sigma(P)$ is the standard deviation of P,
c_1^*, c_2^* are the 2.5% and 97.5% quantiles of $\frac{T(P^*)-T(\hat{P})}{\sigma(P^*)}$,
and P^* denotes the bootstrapped samples, X^*.

To construct an interval for $\mathbb{E}(X)$, we set $T(X) = \frac{1}{n}\sum_{i=1}^n X_i$. Using Equation (3), this gives the interval:

$$(1.0761, 1.3233)$$

To construct an interval for $\mathbb{E}(X^2)$, we set $T(X) = \frac{1}{n}\sum_{i=1}^n X_i^2$. Once again using Equation (3), this gives the interval:

$$(1.7209, 2.3142)$$

4.2 Part b

We wish to construct a permutation test to test the hypothesis:

$$H_0 : X, Y \text{ are independent} \quad H_A : X, Y \text{ are dependent}$$

We are asked to do this by considering the Spearman correlation coefficient. Consider the result that X,Y are independent iff their Spearman correlation coefficient, $\hat{\rho}$, is 0. We can use this to re-state our null hypothesis as:

$$H_0 : \hat{\rho} = 0 \quad H_A : \hat{\rho} \neq 0$$

We can test this null hypothesis by constructing bootstrapped samples of Y. We denote these samples as $(Y_1^*, \dots, Y_n^*)$, and calculate ρ_i as the Spearman Correlation coefficient for (X, Y_i^*) $\forall i \in \{1, \dots, 10000\}$. We construct a p-value for our hypothesis test using:

$$p = \frac{1}{10000}\sum_{i=1}^{10000} \mathbb{1}\{\rho_i \geq \hat{\rho}\}$$

This results in a p-value of 0.0003. Therefore, we reject our null hypothesis at the 95% level of significance and conclude X and Y are not independent.