1. [5 points] Consider a linear congruential pseudorandom number generator, with parameters \( m = 100, \ a = 11, \) and \( b = 3. \) Suppose the initial seed value is \( x_0 = 1. \) Compute the first two values \( U_1 \) and \( U_2 \) that the generator will produce.

**Solution:** Here \( x_1 = ax_0 + b = 11(1) + 3 = 14. \) And, \( x_2 = ax_1 + b = 11(14) + 3 = 157 \equiv 57 \pmod{m}. \) So, \( U_1 = x_1/m = 0.14, \) and \( U_2 = x_2/m = 0.57. \)

2. [5 points] Let \( g(y) = e^y(1 + \sin(y))1_{[0,8]}(y), \) let \( \pi(y) = cg(y) \) be a probability density (with unknown normalising constant \( c \)), and let \( h(y) = y^2. \) Suppose we simulate two samples \( x_1, x_2 \sim \text{Exponential}(3), \) and observe that \( x_1 = 2 \) and \( x_2 = 6. \) Compute the Importance Sampler estimate of \( \mathbb{E}_{\pi}(h) \) based on these two samples.

**Solution:** Here the \( x_i \) samples are from the density \( f(x) = 3e^{-3x} \) \((x > 0)\), so the Importance Sampler estimate is

\[
\frac{\sum_{i=1}^{2} h(x_i) g(x_i) / f(x_i)}{\sum_{i=1}^{2} g(x_i) / f(x_i)} = \frac{(h(2)g(2)/f(2)) + (h(6)g(6)/f(6))}{(g(2)/f(2)) + (g(6)/f(6))} = \frac{2^2e^2(1 + \sin(2)) / 3e^{-3(2)} + 6^2e^6(1 + \sin(6)) / 3e^{-3(6)}}{(e^2(1 + \sin(2)) / 3e^{-3(2)}) + (e^6(1 + \sin(6)) / 3e^{-3(6)})} = \frac{4e^8(1 + \sin(2)) + 36e^{24}(1 + \sin(6))}{e^8(1 + \sin(2)) + e^{24}(1 + \sin(6))} = \frac{4(1 + \sin(2)) + 36e^{16}(1 + \sin(6))}{(1 + \sin(2)) + e^{16}(1 + \sin(6))}. \]

3. Consider the statistical model that \( Y_i \sim \text{Exponential}(a) \) for \( i = 1, 2, 3 \) (conditionally independently), where \( a \in \mathbb{R} \) is unknown. Let \( a \) have prior density given by \( f(a) \propto (1 + a^2)1_{[1,4]}(a) \). Suppose we then observe the data values \( Y_1 = 2, \ Y_2 = 3.5, \ Y_3 = 6.5. \)

(a) [5 points] Conditional on this data, specify (up to a normalising constant) the posterior density of \( a. \)

**Solution:** Here the likelihood function is \( \prod_{i=1}^{3} ae^{-aY_i}, \) so since “posterior equals prior times likelihood”, the posterior density is

\[
\pi(a) = c f(a) \prod_{i=1}^{3} ae^{-aY_i} = c (1+a^2)1_{[1,4]}(a) a^3 e^{-a(2+3.5+6.5)} = c (1+a^2)a^3 e^{-12a}1_{[1,4]}(a). \]
(b) [10 points] Write a complete R program to run an Independence Sampler algorithm with proposal distribution \( N(0, 1) \), to estimate the posterior mean of \( \log(a) \). (Your program should output the final estimate, but you don’t need to bother with standard errors or confidence intervals or acceptance rates or plots or any other “extras”.)

**Solution:**

M = 1100  # run length  
B = 100   # burn-in  
g = function(a) {  
  if ( (a<1) || (a>4) )  
    return(0)  
  else  
    return( (1+a^2) * a^3 * exp(-12*a) )  
}  
alist = rep(0,M)  
a = runif(1,1,4) # overdispersed starting distribution: Uniform[1,4]  
for (i in 1:M) {  
  y = rnorm(1) # proposal  
  A = (g(y) * dnorm(a)) / (g(a) * dnorm(y))  
  U = runif(1)  
  if (U<A)  
    a = y # accept proposal  
  alist[i] = a  
}  
print(mean(log(a[(B+1):M]))) # output the estimate

4. [5 points] Suppose from a Markov chain run with \( M = 200 \) and \( B = 100 \), we obtain the estimates that \( E_\pi(h) \approx 20 \), and \( \text{Var}_\pi(h) \approx 9 \), and \( \text{Corr}_\pi[h(X_n), h(X_{n+1})] \approx 0.6 \), and \( \text{Corr}_\pi[h(X_n), h(X_{n+2})] \approx 0.3 \), and \( \text{Corr}_\pi[h(X_n), h(X_{n+k})] \approx 0 \) for all \( k \geq 3 \). In terms of this, provide an approximate 95% confidence interval for \( E_\pi(h) \), clearly stating any assumptions that you require.

**Solution:**

Here we estimate \( v = \frac{1}{M-B}(\text{Var}_\pi(h)) \text{varfact} \approx \frac{1}{100} 9 \left( 1 + 2 \sum_{k=1}^{\infty} \text{Corr}_\pi[h(X_n), h(X_{n+k})] \right) = 0.09(1 + 2(0.6 + 0.3 + 0)) = 0.09(2.8) = 0.252. \)

Then, we know from class that if a CLT holds (e.g. the chain is geometrically ergodic, and \( E_\pi((h)^{2+\delta}) < \infty \) for some \( \delta > 0 \)), and if the burn-in \( B \) and effective run length \( M - B \) are both sufficiently large, then an approximate 95% confidence interval for \( E_\pi(h) \) is given by:

\[
(e - 1.96\sqrt{v}, e + 1.96\sqrt{v}) = (20 - 1.96\sqrt{0.252}, 20 + 1.96\sqrt{0.252}) \approx (19.016, 20.984).
\]
5. [5 points] Consider a usual Parallel Tempering (MCMCMC) algorithm, with five temperatures \( \tau = 1, 2, 3, 4, 5 \), and with \( \pi_{\tau}(x) = c_{\tau} \left( e^{-|x|^2} \right)^{1/\tau} \) for \( x \in \mathbb{R} \), for some \( c_{\tau} > 0 \). Assume that, as usual, the temperatures for an attempted swap are chosen uniformly at random from \( \{1, 2, 3, 4, 5\} \). Suppose that at time \( n \), the algorithm is in the state \( X_n = (1, 2, 2, 1, 4) \), and then proposes to swap the values corresponding to \( \tau = 3 \) and \( \tau = 5 \). Compute the probability that this proposed swap will be accepted.

**Solution:** We know from class that for symmetrically-chosen temperature pairs \( \tau \) and \( \tau' \), the acceptance probability for a proposed swap is

\[
\min \left( 1, \frac{\pi_{\tau}(X_{n,\tau}) \pi_{\tau'}(X_{n,\tau'})}{\pi_{\tau}(X_{n,\tau}) \pi_{\tau'}(X_{n,\tau'})} \right) = \min \left( 1, \frac{\pi_3(4) \pi_5(2)}{\pi_3(2) \pi_5(4)} \right)
\]

\[
= \min \left( 1, \frac{c_3(4)^{1/3} c_5(2)^{1/5}}{c_3(2)^{1/3} c_5(4)^{1/5}} \right) = \min \left( 1, \frac{(e^{-4^2})^{1/3} (e^{-2^3})^{1/5}}{(e^{-2^4})^{1/3} (e^{-4^2})^{1/5}} \right)
\]

\[
= \min \left( 1, e^{-4(7/3) - (2^7/5) + (2^7/3) + (4^7/5)} \right) = \min \left( 1, e^{-(64/3) - (8/5) + (8/3) + (64/5)} \right)
\]

\[
= \min \left( 1, e^{-64 - 8/(3) + 8/3 - (1/5)} \right) = \min \left( 1, e^{-56(2/15)} \right)
\]

\[
= \min \left( 1, e^{-112/15} \right) = e^{-112/15} \approx 0.0005718.
\]

6. Let \( X = [0, 1] \). Define Markov chain transition probabilities \( P(x, \cdot) \) as follows. If \( x = 1/m \) for some positive integer \( m \), then \( P(x, \cdot) = x^2 \text{Uniform}[0, 1] + (1 - x^2) \delta_{1/(m+1)}(\cdot) \), i.e., it is a mixture distribution (where \( \delta_{1/(m+1)}(\cdot) \) is the point-mass [degenerate] distribution at the point \( 1/(m+1) \)). For all other \( x \in X \), \( P(x, \cdot) = \text{Uniform}[0, 1] \).

(a) [5 points] Does this chain have a stationarity distribution \( \Pi \), and if so what is it?

**Solution:** Yes, \( \Pi = \text{Uniform}[0, 1] \) is stationary, since if we choose \( X_0 \sim \text{Uniform}[0, 1] \), then \( P(X_0 = 1/m \text{ for some positive integer } m) = 0 \), so

\[
P(X_1 \in S) = \int_{x \in X} P(x, S) \Pi(dx) = \int_{x \in X} (\text{Uniform}[0, 1])(S) \Pi(dx)
\]

\[
= (\text{Uniform}[0, 1])(S) = \Pi(S)
\]

as required.

(b) [5 points] Determine (with explanation) the set \( G \subseteq X \) of states \( x \in X \), such that \( \lim_{n \to \infty} P^n(x, S) = \Pi(S) \) for all (measurable) \( S \subseteq X \) if and only if \( x \in G \).

**Solution:** If \( x \notin \{1/2, 1/3, 1/4, \ldots\} \), then \( P^n(x, \cdot) = \Pi(\cdot) \) for all \( n \geq 1 \), so indeed \( \lim_{n \to \infty} P^n(x, S) = \Pi(S) \) for all (measurable) \( S \subseteq X \).
However, if $x = 1/m$ for some integer $m \geq 2$, then $P(X_n = 1/(m + n)$ for all $n \geq 1 \mid X_0 = x) = \prod_{n=0}^{\infty} (1 - (m+n)^{-2}) \geq \prod_{j=2}^{\infty} (1 - j^{-2}) = \exp \left( \sum_{j=2}^{\infty} \log(1 - j^{-2}) \right) \geq \exp \left( - \sum_{j=2}^{\infty} (j^{-2} + \frac{1}{2} j^{-4}) \right) \equiv e^{-K} > 0$, where $K = \sum_{j=2}^{\infty} (j^{-2} + \frac{1}{2} j^{-4}) < \infty$. (We have used that for $|y| < 1$, $\log(1+y) = y - (y^2/2) + (y^3/3) - (y^4/4) + \ldots \geq y - (y^2/2)$.

Alternatively, simply note that $\prod_{j=2}^{\infty} (1 - j^{-2}) > 0$ since $(1 - j^{-2}) > 0$ for all $2 \leq j < \infty$, and $\sum_{j=2}^{\infty} j^{-2} < \infty$.) In particular, if $S = \{1/2, 1/3, \ldots\}$, then $P^n(x, S) \geq P^n(x, \{1/(m+n)\}) \geq e^{-K} \neq 0 = \Pi(S)$, so we do not have $\lim_{n \to \infty} P^n(x, S) = \Pi(S)$ for all (measurable) $S \subseteq X$.

So, $G = X \setminus \{1/2, 1/3, 1/4, \ldots\} = \{x \in X : x \neq 1/m$ for any integer $m \geq 2\}$. 

4