Solutions to Sample Midterm Questions

1. Parameter fitting.

(a) Density and likelihood function

(i) The density is 
\[ f_X(x) = \frac{d}{dx} F_X(x) = \frac{a}{b} \left( \frac{x}{b} \right)^{-(a+1)} \] for \( x \geq b \) and 
\[ f_X(x) = 0 \] for \( x < b \)

(ii) The likelihood function is 
\[ L_n(a,b; X_1, \ldots, X_n) = \prod_{i=1}^{n} f_X(X_i) = \prod_{i=1}^{n} I(b \leq X_i) \times a^n b^\sum_{i=1}^{n} X_i^{-(a+1)} \]

where \( I \) is an indicator function.

(b) Since \( L_n \) is initially increasing in \( b \), choose \( b \) as large as possible: 
\[ \hat{b} = X_{(1)} = \min_i X_i. \]

To choose \( \hat{a} \), maximize the function 
\[ h(a) = a^n b^\sum_{i=1}^{n} X_i^{-(a+1)} \]

Equivalently, maximize 
\[ \tilde{h}(a) = \log h(a) = n \log a + n a \log \hat{b} - (a + 1) \sum_{i=1}^{n} \log X_i. \]

\[ \frac{d}{da} \tilde{h}(a) = \frac{n}{a} + n \log \hat{b} - \sum_{i=1}^{n} \log X_i = 0 \]

implies 
\[ \hat{a} = \left[ \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{X_i}{X_{(1)}} \right) \right]^{-1}. \]

(c) Given \( \hat{b} \) we have
\[ E[X] = \int_{b}^{\infty} x f_X(x) \, dx = \int_{b}^{\infty} ab^\frac{x}{b} x^{-(a+1)} \, dx = \frac{ab^\frac{x}{b} x^{-(a+1)} - \frac{a}{b} x^{-(a+1)}}{-1} \bigg|_{b}^{\infty} = \frac{a \hat{b}}{a - 1}. \]

Substitute the sample mean \( \bar{X}_n \) for \( E[X] \) and solve the equation
\[ \bar{X}_n = \frac{a \hat{b}}{a - 1} \]

to get 
\[ \hat{a} = \frac{\bar{X}_n}{\bar{X}_n - b}. \]

2. Random variate generation

(a) The picture is as follows.

Using the basic version of the acceptance-rejection method, we obtain the following algorithm:
Algorithm:
1. Generate uniform random numbers $U_1, U_2$
2. Set $Z_1 = bU_1$ and $Z_2 = U_2 / (1 - e^{-b})$
3. If $Z_2 \leq e^{-Z_1}/(1 - e^{-b})$, then return $Z_1$, else go to Step 1.

(b) A much more efficient inversion algorithm is as follows, based on the fact that $F_X(x) = 1 - e^{-x}$.

Algorithm:
1. Generate a uniform random number $U$
2. Return $X = -\ln(1 - (1 - e^{-b})U)$

3. Linear congruential generators

a) Starting with $x_0 = 1$, we have the following results

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x_i$</th>
<th>$3i \mod 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

By inspection of the second column, we see that all values between 1 and 6 occur, so the generator has full period. Inspection of the third column shows that 3 is a primitive element modulo 7, which also proves that the generator has full cycle since 7 is prime.

b) From Lecture 6, Slide 8, the RANDU generator is an MCG with modulus $2^{31}$. By the Claim on Slide 10, the period is at most $2^{4-2} = 2^{2-2} = 2^4 = 16$.

4. Bayesian estimation. Write $X_{(n)} = \max_{1 \leq i \leq n} X_i$, and note that $X_i \geq 0, \forall i$. Then the likelihood is given by $f(X_{(a)} | a) = \prod_{i=1}^{n} [I(X_i \leq a) / a] = I(X_{(a)} \leq a) / a^n$, where $I(A)$ denotes the indicator of event $A$ and $X_{(a)} = X_1, \ldots, X_n$. The prior is $f(a) = I(0 \leq a \leq b) / b$. By Bayes’ Theorem, the posterior density of $a$ is

$$f(a | X_{(a)}) = \frac{f(X_{(a)} | a)f(a)}{\int f(X_{(a)} | a)f(a)da} = \frac{I(X_{(a)} \leq a)I(0 \leq a \leq b)a^{-b}b^{b-1}}{\int I(X_{(a)} \leq a)I(a \leq b)a^{-b}b^{b-1}da} = \frac{I(X_{(a)} \leq a \leq b)a^{-b}}{\int I(X_{(a)} \leq a \leq b)a^{-b}da} = \frac{I(X_{(a)} \leq a \leq b)a^{-b}}{(X_{(a)}^{-a} - b^{-a})/(n-1)}.$$
Thus the posterior mean is \( \hat{a} = \int_{a_1}^{a_n} \frac{a^{\frac{1}{1-a}}}{a^{\frac{1}{1-a}} - b^{\frac{1}{1-a}} / (n-1)} \)

\( \int_{X_{(s)}}^{b} a^{\frac{1}{1-a}} da = \frac{(X_{(s)} - b^{\frac{1}{1-a}}) / (n-1)}{(X_{(s)} - b^{\frac{1}{1-a}}) / (n-1)} \).