# Math 447 Assignment 4

*Jonathan Pearce 260672004*

*3/28/2018*

## 5.7

Since the proposal distribution is uniform, $T_{ij} = T_{ji}\ \forall i, j$, and we have,

$$\pi_k = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad k = 0, 1, 2...$$

Therefore the accpetance function is reduced to,

$$a(i,j) = min\Big(1, \frac{\pi_j T_{ji}}{\pi_i T_{ij}}\Big) = min\Big(1, \frac{\frac{n!}{j!(n-j)!} p^j (1-p)^{n-j}}{\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}}\Big) = min\Big(1, \frac{i!(n-i)!}{j!(n-j)!} p^{j-i} (1-p)^{i-j}\Big)$$

Algorithm Procedure:

1. Randomly select an initial state $X_0 \in \{0, 1, 2, 3....n\}$
2. for N=1,2,3,. . . ,k
   - Propose to move from $X_{N-1} = i$, to $X_N = j$. Where j is selected from the uniform proposal distribution $T$ such that $i \neq j$.
   - Compute $a(i, j)$
   - Let $U$ be uniformly distributed on $[0, 1]$. If $U \leq a(i, j)$, $X_N = j$. Otherwise, $X_N = i$.

## 5.17

### a & b

Conditional distribution of $X$ given $N = n$ is proportional to $e^{-3x}x^n$ for $x > 0$. Which gives a gamma distribution with shape parameter $n + 1$ and rate parameter 3.

Conditional distribution of $N$ given $X = x$ is proportional to $\frac{e^{-x}x^n}{n!}$ for $n = 0, 1, 2, ....$ Which gives a poisson distribution with parameter $x$.

```
set.seed(107)

iCount = 0
iiSum = 0
trials = 1000000

sim = matrix(rep(0,2*trials),ncol=2)
sim[1,] = c(1,0) # (x,n)

for (i in 2:trials){
  sim[i,1] = rgamma(1,(sim[i-1,2] + 1),3) #sample x from gamma
  sim[i,2] = rpois(1,sim[i,1]) #sample n from poisson

  if (sim[i,1]^2 < sim[i,2]){ #count for question i)
    iCount = iCount + 1
```

```
  }
  iiSum = iiSum + (sim[i,1]*sim[i,2]) #sum of X*N for ii)

}
iAns = iCount/trials
iiAns = iiSum/trials
iAns
```

```
## [1] 0.263882
```

```
iiAns
```

```
## [1] 0.5001157
```

Therefore $P(X^2 < N) \approx 0.263882$ and $E(XN) \approx 0.5001157$

# 5.19

```
set.seed(108)

n = 50
p = 1/4.0
trials = 1000000

i = sample(0:50, 1)

#simValues <- matrix(rep(0,1*trials),ncol=1)
#simValues[1,1] = i

count = 0
if(i >= 10 && i <= 15){
    count = 1
  }

for (k in 2:trials){
  j = sample(0:50, 1) #generate proposal state
  #ensure j =/= i
  while (i==j){
    j = sample(0:50, 1)
  }

  #calculate a
  a = ((factorial(i)*factorial(n-i))/(factorial(j)*factorial(n-j)))*(p^(j-i))*((1-p)^(i-j))
  if(a > 1){
    a = 1
  }

  #decide whether to accept proposal
  u = runif(1,0,1)
  if(u < a){
    i = j
  }
  #simValues[k,1] = i
```

```
  if(i >= 10 && i <= 15){
    count = count + 1
  }

}
ans = count/trials
ans
```

```
## [1] 0.671691
```

Our implementation calculated $P(10 \leq X \leq 15) = 0.671691$. The true value of this probability is 0.6732. The percent error of our approximation is 0.2242%.

## 6.3

### a.

This is equivilant to calculating the probability that the first arrival time is greater than 30 seconds (0.5 minutes)

$$P(X > 0.5) = e^{-5*\frac{1}{2}} \approx 0.08208$$

### b.

$$P(N_1 = 4, N_2 - N_1 = 6)$$

Since the time intervals are disjoint

$$= P(N_1 = 4)P(N_2 - N_1 = 6)$$

$$= P(N_1 = 4)P(N_1 = 6)$$

$$= \frac{e^{-5*1} * (5*1)^4}{4!} * \frac{e^{-5*1} * (5*1)^6}{6!}$$
$$= 0.025657$$

### c.

$$P(N_5 = 25, N_1 = 6)$$
$$= P(N_5 - N_1 = 19, N_1 = 6)$$

Since the time intervals are disjoint

$$= P(N_5 - N_1 = 19)P(N_1 = 6)$$

$$= P(N_4 = 19)P(N_1 = 6)$$

$$= \frac{e^{-5*4} * (5*4)^{19}}{19!} * \frac{e^{-5*1} * (5*1)^6}{6!}$$
$$= 0.01299$$

## 6.9

First,

$$\begin{aligned} P(X > x) &= 1 - P(X \leq x) \\ &= 1 - (1 - (1-p)^x) \\ &= (1-p)^x \end{aligned}$$

Now,

$$\begin{aligned} & P(X > s + t \mid X > s) \\ &= \frac{P(X > s+t, X > s)}{P(X > s)} \\ &= \frac{P(X > s+t)}{P(X > s)} \\ &= \frac{(1-p)^{s+t}}{(1-p)^s} \\ &= (1-p)^t \\ &= P(X > t) \end{aligned}$$

## 6.28

Lets examine the first time period, i.e. the first interval that a tornado could hit the region. Let $T_t$ be the number of tornadoes by time $t$. We have,

$$P(T_1 = 0) = e^{-2}$$

$$P(T_1 > 0) = 1 - e^{-2}$$

Let $C_1$ be the number claims filed because of the tornadoes from the first time period.

$$P(C_1 = k) = \begin{cases} e^{-2} + (1 - e^{-2})\frac{30^k e^{-30}}{k!}, & \text{if } k = 0 \\ (1 - e^{-2})\frac{30^k e^{-30}}{k!}, & \text{if } k \neq 0 \end{cases}$$

There will only be claims filed when $T_1 > 1$. Therefore,

$$\begin{aligned} E[C_1] &= \sum_{k=0}^{\infty} kP(C_1 = k) \\ &= \sum_{k=1}^{\infty} kP(C_1 = k) \\ &= \sum_{k=1}^{\infty} k(1 - e^{-2})\frac{30^k e^{-30}}{k!} \\ &= 30(1 - e^{-2})e^{-30} \sum_{k=1}^{\infty} \frac{30^{k-1}}{(k-1)!} \\ &= 30(1 - e^{-2})e^{-30} \sum_{k=0}^{\infty} \frac{30^k}{k!} \\ &= 30(1 - e^{-2})e^{-30}e^{30} \end{aligned}$$

$$= 30(1 - e^{-2})$$

$$\begin{aligned}
E[C_1^2] &= \sum_{k=0}^{\infty} k^2 P(C_1 = k) \\
&= \sum_{k=1}^{\infty} k^2 P(C_1 = k) \\
&= \sum_{k=1}^{\infty} k^2 (1 - e^{-2}) \frac{30^k e^{-30}}{k!} \\
&= 30(1 - e^{-2}) e^{-30} \sum_{k=1}^{\infty} k \frac{30^{k-1}}{(k-1)!} \\
&= 30(1 - e^{-2}) e^{-30} \sum_{k=1}^{\infty} ((k-1) + 1) \frac{30^{k-1}}{(k-1)!} \\
&= 30(1 - e^{-2}) e^{-30} \sum_{k=1}^{\infty} (k-1) \frac{30^{k-1}}{(k-1)!} + \frac{30^{k-1}}{(k-1)!} \\
&= 30(1 - e^{-2}) e^{-30} \Big( 30 \sum_{k=1}^{\infty} \frac{30^{k-2}}{(k-2)!} + \sum_{k=1}^{\infty} \frac{30^{k-1}}{(k-1)!} \Big) \\
&= 30(1 - e^{-2}) e^{-30} \left( 30 \Big( \frac{30^{-1}}{(-1)!} + \sum_{k=2}^{\infty} \frac{30^{k-2}}{(k-2)!} \Big) + e^{-30} \right) \\
&= 30(1 - e^{-2}) e^{-30} \left( 30 \Big( \frac{30^{-1}}{(-1)!} + e^{-30} \Big) + e^{-30} \right) \\
&= 30(1 - e^{-2}) e^{-30} \left( -1 + 31 e^{-30} \right) \\
&= -30(1 - e^{-2}) e^{-30} + 31 * 30(1 - e^{-2}) e^{-30} e^{-30} \\
&= 31 * 30(1 - e^{-2})
\end{aligned}$$

The Variance follows directly

$$\begin{aligned}
\Rightarrow Var(C_1) &= E[C_1^2] - (E[C_1])^2 \\
&= 31 * 30(1 - e^{-2}) - (30(1 - e^{-2}))^2 \\
&= 30(1 - e^{-2})(31 - 30(1 - e^{-2}))
\end{aligned}$$

The number of tornadoes that hit the region each time interval is indepedent of the other time intervals. Therefore, we have that the number of claims filed due to tornadoes in each time interval is independent of claims filed due to tornadoes from different time intervals. Therefore the total number of claims filed by time t is simply a sum of t independent single intervals, thus

$$E[C_t] = E[C_1] * t = 30(1 - e^{-2})t$$

$$Var(C_t) = Var(C_1) * t = 30(1 - e^{-2})(31 - 30(1 - e^{-2}))t$$