Name: ____________________________

Directions: **Work only on this sheet** (on both sides, if needed). MAKE SURE TO COPY YOUR ANSWERS TO A SEPARATE SHEET FOR SENDING ME AN ELECTRONIC COPY LATER.

Important note: Remember that in problems calling for R code, you are allowed to use any built-in R function, e.g. **choose()**, **pnorm()**, etc.

1. Consider the class enrollment size example, starting on p.97. Suppose the distribution of enrollment size is Poisson, rather than approximate normal. Assume the mean is again 28.8.

(a) (20) Find Var(N).

(b) (20) Find $F_N(26)$.

(c) (15) Find $P(N \geq 30|N \geq 25)$.

2. Consider the network intrusion example on p.97.

(a) (15) Let $Y = Z^2$. Name the parametric family of densities that Y's density belongs to, including the parameter values, if any.

(b) (15) Let G denote the indicator random variable for the event $X \geq 535$. Find Var(G).

3. (15) Suppose R didn't include the **sample()** function. We could use the code below instead. Here's an example of usage:

```
> x <- samp(c(1,6,8),1000,c(0.2,0.5,0.3))
> sum(x==1)
[1] 224
> sum(x==6)
[1] 495
> sum(x==8)
[1] 281
```

Here we generate 1000 numbers from 1,2,3, with probability 0.2, 0.5 and 0.2, respectively, and then count the numbers of 1s, 6s and 8s we get.

The built-in R function **cumsum()** finds cumulative sums, e.g.

```
> cumsum(c(3,8,1))
[1]  3 11 12
```

```
# what we'd do if there were no sample()
# ftn in R; sample n items (with
# replacement) from the vector nums, with
# probabilities given by prob;
# the vectors nums and prob must have
# the same length (not checked here);
# not claimed efficient
samp <- function(nums,n,prob) {
   samps <- vector(length=n)
   cumulprob <- cumsum(prob)
   for (i in 1:n)
      samps[i] <- sample_one_item(nums,cumulprob)
   return(samps)
}

sample_one_item <- function(nums,cumulprob) {
   u <- runif(1)
   lc <- length(cumulprob)
   for (j in 1:(lc-1)) {
      if (BLANKa) BLANKb
   }
   BLANKc
}
```

Solutions:

1.a Since N has a Poisson distribution, Var(N) = E(N) = 28.8.

1.b For a Poisson random variable M, $\lambda = EM$, so answer is

```
ppois(26,28.8)
```

1.c (4.50) still holds, and evaluates to

```
(1 - ppois(29,28.8)) / (1 - ppois(24,28.8))
```

2.a Chi-square, 1 degree of freedom.

2.b From Section 3.6:

$$Var(G) = P(G=1) \cdot [1 - P(G=1)] = 0.01 \cdot 0.99 \tag{1}$$

3.

```
# what we'd do if there were no sample() ftn in R; sample n items (with
# replacement) from the vector nums, with probabilities given by prob;
# the vectors nums and prob must have the same length (not checked here);
# not claimed efficient
samp <- function(nums,n,prob) {
   samps <- vector(length=n)
   cumulprob <- cumsum(prob)
   for (i in 1:n)
      samps[i] <- sample_one_item(nums,cumulprob)
   return(samps)
}

sample_one_item <- function(nums,cumulprob) {
   u <- runif(1)
   lc <- length(cumulprob)
   for (j in 1:(lc-1)) {
      if (u < cumulprob[j]) return(nums[j])
   }
   return(nums[lc])
}
```