NYCCT MAT 2572 Halleck Fall 2017 Practice exam 3 solutions

The exam will consist of problems similar to the following 5 problems (20% each). You may use your calculator and/or your computer (Excel, Maple, MATLAB, R) as an aid. You are encouraged to bring in a formula sheet and instructions for each problem and even a template for software use. On the exam, I will keep the same numbering to avoid confusion.

1. On-the-job injuries in a textile mill severe enough to lead a worker to ask for disability leave occur at the rate of 0.1 per day. We model using the Poisson distribution.
   a. What 2 assumptions do we make? **We are assuming that each injury is independent of others. Also we are assuming that the accident rate does not change.**
   What may make these assumptions unrealistic for our situation? **In one accident, several workers may be injured. The rate of accidents may change depending on factors such as night-before parties or end-of-work-week fatigue.**
   b. What is the probability that two injuries will occur during the next (six-day) workweek?

$$\lambda = \mathbf{6 * .1 = .6\ \&\ } \boldsymbol{p(k) = \lambda^k e^{-\lambda}/k!},$$
$$\textbf{so } \boldsymbol{p(2)} = \frac{.6^2 e^{-.6}}{2!} = \frac{\mathbf{0.6^2 EXP(-0.6)}}{\mathbf{2}} = \mathbf{0.098786 \approx 10.0\%}$$

   c. The probability that four injuries will occur over the next two workweeks is not the square of your answer to part (a). Explain why not and find it.

The square of the probability from b would find the chance that exactly 2 injuries happen in week 1 and exactly 2 injuries happen in week 2. However, 4 injuries in 2 weeks could be distributed in many different ways, e.g., 0 and 4 or 1 and 3.

$$\lambda = \mathbf{12 * .1 = 1.2\ \&\ } \boldsymbol{p(4)} = \frac{\mathbf{1.2^4 e^{-1.2}}}{\mathbf{4!}} = \frac{\mathbf{1.2^2 EXP(-1.2)}}{\mathbf{24}} = \mathbf{0.026023 \approx 2.6\%}$$

2. Recently married, a young couple plans to continue having children until they have their first girl. Our "experiment" is to count the number of boys born until that girl is born. We model using the geometric distribution with p = 1/2.
   a. What 2 assumptions do we make?
      i. **Mother can have unlimited number of children.**
      ii. **One child is born each time.**
      iii. **Chance of a girl is 50%**
   b. What may make these assumptions unrealistic for our situation?
      i. A woman can have a limited number of children within her lifetime, perhaps 20 or 30, assuming a single child is born each time. The world record is 69, but that was from a woman who had multiple children births many times.
      ii. Identical twins have the same gender, so the gender of each child is not always independent of the previous child.
      iii. Chance of a girl child is not exactly .5; it is believed to be ~49%: http://www.thebump.com/a/more-baby-boys-than-girls
   c. On average, how many boys will be born? What is the couple's expected family size?

   Using the formula q/p, we get 1 for the number boys expected. So family size is 1 boy + 1 girl + 2 parents = 4.

   d. What is the variance for the number of boys? For the family size? Are they different? If not, why not?

   Variance for the number of boys is $\frac{q}{p^2} = \frac{\frac{1}{2}}{\left(\frac{1}{2}\right)^2} = \mathbf{2}$**. Family size = #boys + 3. Thus, family size variance will be the same. Variance is not affected by horizontal shifting.**

3. A machine has a 1% probability of producing a defective item. Each day, the machine is run until a defective item is produced and then it undergoes extensive repair which requires the rest of the day. We count the number of usable items produced in one 5-day workweek and model using negative binomial distribution.
   a. What is the average number of usable items that will be produced?

      r=5 and $\boldsymbol{E(X) = \frac{rq}{p} = \frac{5*.99}{.01} = 495}$

   b. What is the standard deviation?

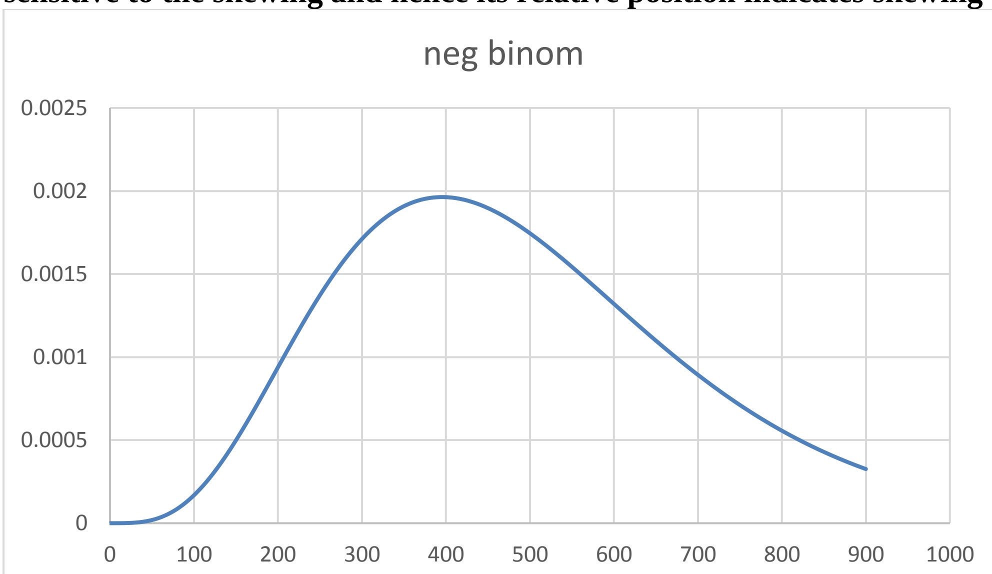
$$\boldsymbol{V(X) = \frac{rq}{p^2} = \frac{5*.99}{.01^2} = 49500 \text{ and } \sigma \approx 223}$$

   c. What is the probability that a machine will produce 500 or more usable items?

$$\boldsymbol{P(X \geq 500) = 1 - P(X \leq 499) = 1 - \sum_{k=0}^{499}\binom{5+k-1}{5-1}0.01^5 0.99^k = 1 - NEGBINOM.DIST(499, 5, 0.01, TRUE) \approx 43.3\%}$$

   d. Draw an approximate graph of the distribution, and find approximation for the mode. Also use Excel to find an approximation for the median (see excel file). Comment on their relative positions and the skewing that is present.

      See Excel file: Mode is ~400, median is ~ 465 and mean is 495. Note that mean is most sensitive to the skewing and hence its relative position indicates skewing to the right.

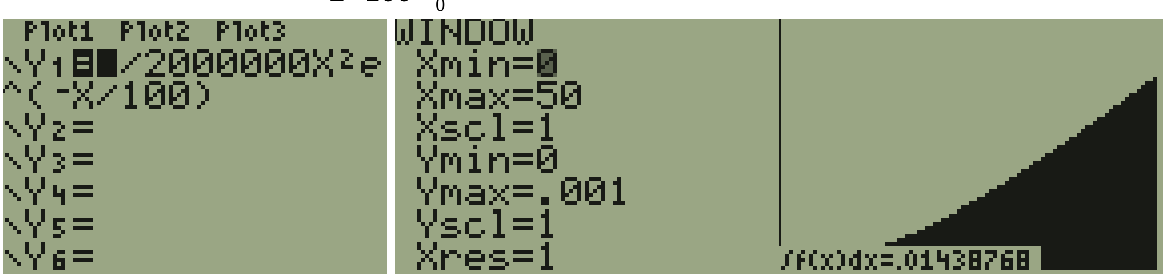


4. The next generation of space shuttle will include three fuel pumps —one active, the other 2 in reserve. If the primary pump malfunctions, a second is automatically brought on line. A mission will require that fuel be pumped for 50 hours. The average lifespan of a pump is 100 hours. We model using the gamma distribution.
   a. If the pumps are allowed to go until all 3 failed, on average how many hours will that be?

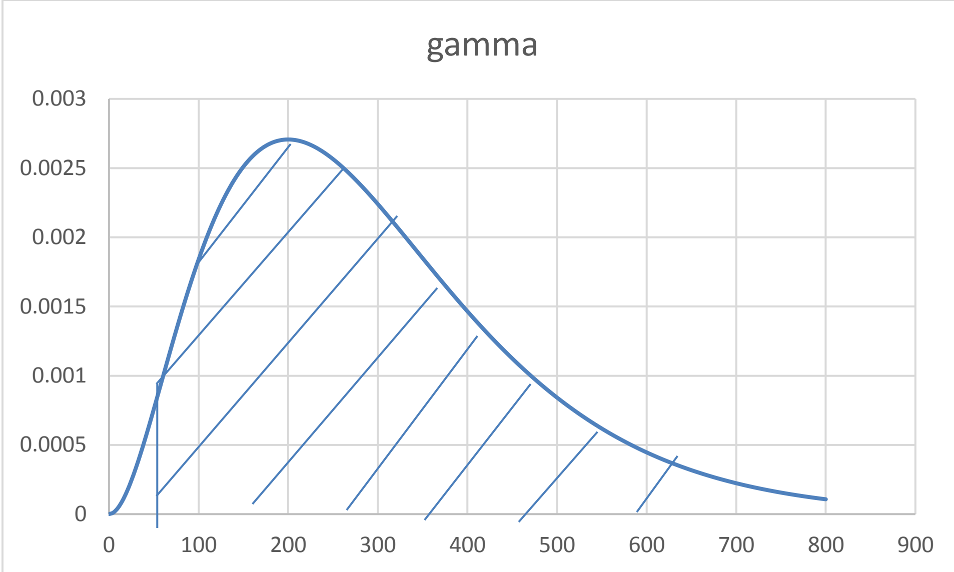
      r=3, $\boldsymbol{\lambda = \frac{1}{100}}$ **and** $\boldsymbol{E(Y) = \frac{r}{\lambda} = \frac{3}{.01} = 300}$ **hours**

   b. What are the chances that the fuel pump system will not remain functioning for the mission?

To calculate that the system will remain functioning for 50 hours, we use complement and an integral of the pdf: $1 - \frac{1}{2*100^3}\int_0^{50} y^2 e^{-y/100}\,dy = 1 - 0.0144 = 0.9856 \sim 98.6\%$



   c. Find the median for the fuel system's life. Median is ~268 (from Excel table)
   d. Draw an approximate graph of the distribution, marking off the mode (exact), mean (a.) and median (use excel). Also, shade the portion of the graph corresponding to b.



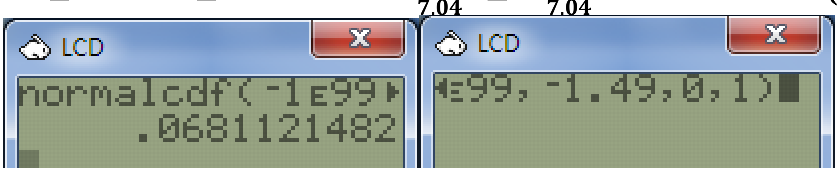
See Excel file: Mode is 200=(r-1)/λ, median is ~ 268 and mean is 300. Note that mean is most sensitive to skewing and hence its relative position indicates skewing to right.

5. There are 2 candidates in a high school student government election, 55% of the students favor the incumbent in her bid for re-election. If 200 students vote, approximate the probability that
   a. the race ends in a tie;
   b. the challenger scores an upset victory using the binomial distribution.
   c. the challenger scores an upset victory using the normal distribution w/ continuity correction.
   d. Draw a picture showing the normal distribution (-3<Z<3) and shade the region corresponding to part c. Provide both standard and nonstandard labels.

We model this problem as a coin flip problem with p =.55, n=200. [This is actually an ill-defined problem, because we don't have information about how likely the supporters of each candidate are to vote. In real life, it would not be the support among the population that counts, but the support among those who are likely to vote.]

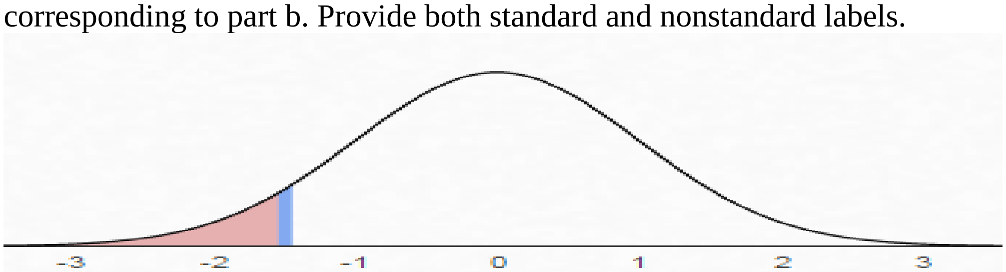
a. the race ends in a tie;
   binomial: k = 100 => P(k) = 200 nCr 100 $.55^{100}.45^{100}$ = 0.02063 ≈ 2.1%
b. the challenger scores an upset victory using the binomial distribution.
   =BINOM.DIST(99, 200, 0.55, TRUE) = 0.06807525 ≈ 6.8%
c. the challenger scores an upset victory (use the normal distribution w/ continuity correction).
   μ=np=110 $\boldsymbol{\sigma = \sqrt{npq} = \sqrt{200*.55*.45} \approx 7.04}$

$\boldsymbol{X < 100 \to X \le 99 \to X \le 99.5 \to Z = \frac{X-110}{7.04} \le \frac{99.5-110}{7.04} = -1.49}$ **so P(X<100)=.068=6.8%**



[Which closely agrees with the exact, binomial calculation (n is high enough for the middling p).]

d. Draw a picture showing the normal distribution (-3<Z<3) and shade the region corresponding to part b. Provide both standard and nonstandard labels.



89 96 103 110 117 124 131

e.