ISYE 6501
Homework 10

October 25, 2019

Question 13.1

1. **Binomial**. The binomial distribution has PDF $f_x(x) = \binom{n}{x} p^x (1-p)^{n-x}$ where n is the number of independent trials, x is the target number of successes, and p the probability of a success. Let's suppose that p is the probability that a MSA student finds a bug in the code of a Python script; the MSA students are assigned the task independently of each other and cannot communicate. The next day, the instructor calls out students at random and asks them to show their work. The task is hard and, based upon estimates from prior cohorts, the probability of a success is $p = 0.1$. The probability of the instructor calling out, say, 5 students who successfully debugged the code within the first 15 calls is given by $f_x(5) = \binom{15}{5} 0.1^5 (0.9)^{10}$.
2. **Geometric**. Sometimes, we are interested in the number of failures before success is achieved on a number of independent trials. This is the case where we use negative binomial distributions, the most common being the geometric distribution. A geometric distribution has PDF $f_x(x) = (1-p)^{x-1} p$ where p represents the probability of an independent event happening, $(1-p)$ is its complement, and x are the number of successive trials considered. It is easy to see that this distribution can model many real-world situations. Wrapping up the previous example of our MSA instructor, p still represents the probability of a debug. When the instructor calls out students at random, $f_x(x=0)$ is the probability that the first student finds the bug, $f_x(x=1)$ is the probability that the second student finds the bug (i.e. one failure), $f_x(x=2)$ is the probability that the third student finds the bug (i.e. two failures), and so forth.
3. **Poisson**. The Poisson, exponential, and Weibull distribution are widely used in reliability engineering to model the life of a component or a system. Reliability engineering is survival analysis of machines and components instead of humans, and reliability R at a time t is defined as the complement of the CDF of our component (i.e. $R(t) = 1 - F_t(t)$). In other words, reliability subtracts out the cumulative probability of a failure up to time t and therefore informs us on how "reliable" our system or our component is. The Poisson distribution informs us on the probability of n events happening within a given a timeframe t. The expression for the Poisson's PDF is $f_t(t, \lambda, n) = \frac{(\lambda t)^x e^{-\lambda t}}{n!}$ where λ is the failure rate of a component. This formula can be used to model, say, the probability of observing 5 failures of a sensor with a failure rate of $\lambda = 0.003$ over 1 year. This formula is helpful when we have

to figure out how many spare sensors to inventor and we want to achieve a given reliability.

4. **Exponential**. The exponential distribution with PDF $f_t(t) = \lambda e^{-\lambda t}$ is linked to the Poisson distribution in that it focuses on the time to failure rather than number of failures. Before, we used the Poisson distribution to model the number of failures of a sensor in a year. Implicitly, the Poisson distribution assumes that the time between each failure is exponential and captured by the exponential distribution. Therefore, it tells us the probability that an individual sensor with failure rate of $\lambda = 0.003$ will fail in a given time-frame.
5. **Weibull** Lastly, the Weibull distribution is also widely implemented in the study of reliability, when the failure rate is not constant over time and is a function of two parameters known as the shape parameter (β, but notation might vary) and location parameter (θ). Clearly, this distribution captures more of the real-world processes that we observe in manufacturing because it allows for the failure rate to be one decreasing over time ($0 < \beta < 1$) or one increasing over time ($\beta > 1$). When $\beta = 0$, the failure rate becomes constant and the Weibull distribution collapses on the exponential distribution. In class, we gave many examples of processes that experience decreasing/increasing failure rate. However, the same one manufacturing process can be modelled as a decreasing failure rate or increasing failure rate depending on the stage of production. Typically, we observe many failures at the beginning because of high infant of mortality of the weakest parts. This is the reason why, say, microprocessors' manufacturers envision burn-in periods during which components are produced and tested but not sold. The "lemon" microprocessors would fail early in the screening process, and the failure rate would stabilize after some time. This burn-in phase is modelled with a Weibull distribution with $0 < \beta < 1$. Once the microprocessors are sold and assembled to produce computers, they are subject to increasing failure rate due to wear-out ($\beta > 1$).