From last time:

For the geometric distribution,

\[ p(x) = q^{x-1}p, \quad x=1, 2, 3, \ldots \]

\[ E(X) = \frac{1}{p}, \quad V(X) = \frac{1}{p^2} \]

\[ x = \text{trial on which 1st success occurs} \]

---

Negative Binomial or Pascal distribution.

Run Bernoulli trials until the r-th success occurs.

\[ Y = \text{trial on which the r-th success occurs} \]

\[ p(y) = \text{Prob}(\text{r-th success occurs on trial } y) \]

\[ = \text{Prob}(\text{r-1 success in the first } y-1 \text{ trials} \cap \text{success on trial } y) \]

\[ = (y-1) \binom{y-2}{r-1} p^{r-1} q^{y-r} \cdot p \]

\[ = (y-1) \binom{y-2}{r-1} p^r q^{y-r} \quad y = r, r+1, \ldots \]

Find \( \mu_y \) and \( \sigma^2_y \)

If \( y \) has a negative distribution, then

\[ Y = X_1 + X_2 + X_3 + \ldots + X_r \]

Each \( X_i \sim \text{Geom}(p) \)
\[ E[Y] = E[X_1 + \ldots + X_r] = \frac{r}{p} \]
\[ V[Y] = V[X_1 + \ldots + X_r] = \frac{r\sigma^2}{p^2} \]

Other definition of the Negative Binomial:

\[ X = \text{# failures before the r\textsuperscript{th} success} \]
\[ = Y - r \quad (Y = X + r) \]

\[ p(x) = \binom{x+r-1}{r-1} p^r q^x, \quad x = 0, 1, 2, \ldots \]

\[ E[X] = E[Y - r] = \frac{r}{p} - r = \frac{r - rp}{p} = \frac{rp}{p} = \frac{r}{p} \]

\[ V[X] = V[Y - r] = V[Y] = \frac{r\sigma^2}{p^2} \]

\[ \text{Example: Each time we fire our cannon,} \]
\[ \quad \text{we have 20\% chance of hitting our} \]
\[ \quad \text{target. It takes 3 hits to sink} \]
\[ \quad \text{the target.} \]
\[ \quad \text{Find the prob. that we have 7 failures} \]
\[ \quad \text{before we sink the target.} \]

\[ \text{NB}(r = 3, \quad p = .2) \]
\[ P(X=7) = p(7) \]
\[ = \binom{7+3-1}{3-1} \cdot (0.2)^3 \cdot (0.8)^7 \]
\[ = \binom{9}{2} \cdot (0.2)^3 \cdot (0.8)^7 \approx 0.0604 \]

Find \( \mu \) and \( \sigma^2 \)
\[ \mu = \frac{\mu^2}{\sigma} = \frac{3 \cdot (0.9)}{0.2} = 12 \]
\[ \sigma^2 = \frac{\mu^2}{\sigma^2} = \frac{3 \cdot (0.8)}{(0.2)^2} = 60 \]

Find the probability that you see more than 10 failures prior to the 3rd success.

\[ P(X > 10) = p(11) + p(12) + \ldots \]
\[ = 1 - P(X \leq 10) \]
\[ = 1 - \left[ p(0) + p(1) + \ldots + p(10) \right] \]
Another Calculus Problem:

\[ \lim_{n \to \infty} (1 + \frac{x}{n})^n \quad \text{let } y = (1 + \frac{x}{n}) \]

Then \( n \ln(y) = n \ln(1 + \frac{x}{n}) \)

\[ = \lim_{n \to \infty} \frac{\ln(1 + \frac{x}{n})}{\frac{1}{n}} \quad \left( \frac{0}{0 \text{ indeterminant form}} \right) \]

\[ = \lim_{n \to \infty} \frac{\frac{1}{(1 + \frac{x}{n})} \cdot \frac{x}{n^2}}{\frac{1}{n}} \quad \text{by L'Hôpital} \]

\[ = \lim_{n \to \infty} \frac{x}{(1 + \frac{x}{n})} = x \]

\[ \lim_{n \to \infty} n \ln(y) = x \]

\[ \ln(\lim_{n \to \infty} y) = x \]

\[ \lim_{n \to \infty} y = e^x \]
Start with the binomial distribution:

\[ p(n) = \binom{n}{x} p^x q^{n-x} \]

And let \( n \to \infty \) but hold \( np = \text{constant} = \mu \)

\[
\lim_{n \to \infty} p(n) = \lim_{n \to \infty} \frac{n!}{x! (n-x)!} \left( \frac{\mu}{n} \right)^x (1 - \frac{\mu}{n})^{n-x}
\]

\[
= \lim_{n \to \infty} \frac{n!}{x! (n-x)!} \frac{\mu^x}{n^x} \frac{(1 - \frac{\mu}{n})^n}{(1 - \frac{\mu}{n})^x}
\]

\[
= \frac{\mu^x}{x!} \lim_{n \to \infty} \frac{n(n-1) \cdots (n-x+1)}{n^x} \left(1 - \frac{\mu}{n}\right)^x
\]

\[
= \frac{\mu^x}{x!} \lim_{n \to \infty} \left(1 - \frac{\mu}{n}\right)^n \frac{1}{(1 - \frac{\mu}{n})^x}
\]

\[
= \frac{\mu^x}{x!} e^{-\mu}
\]

The Poisson distribution

\[ p(n) = \frac{\mu^x}{x!} e^{-\mu} \quad n = 0, 1, 2, \ldots \]
Check \( \sum_{n=0}^{\infty} p(n) = 1 \)?

\[
\sum_{n=0}^{\infty} \frac{\mu^n}{n!} e^{-\mu} = e^{-\mu} \left[ 1 + \mu + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \ldots \right]
\]

Taylor series for \( e^\mu \)

\( = 1 \checkmark \)

\[ E[X] = \mu \]

\[ V[X] = \lim_{n \to \infty} \text{Var}(n \sim \mu) = \lim_{n \to \infty} \mu (1 - \frac{\mu}{n}) = \mu \]

Poisson process

- Observe a continuous process over a fixed amount of time
- \( X \) counts the # of occurrences of a certain type
- Observations in non-overlapping intervals are independent
- In a very small interval, the probability of seeing an occurrence is proportional to the size of that interval
- In a very small thread, the probability of seeing more than 1 occurrence is not 0.

Example:

A fire station gets an average of 2.83 calls per day. (\(\mu\))

On a given day, find the prob of 4 or fewer calls.

\[
p(n) = \frac{2.83^n}{n!} e^{-2.83}, \quad n = 0, 1, 2, \ldots
\]

\[
P[X \leq 4] = p(0) + p(1) + p(2) + p(3) + p(4)
\]

\[
= e^{-2.83} \left[ 1 + 2.83 + \frac{2.83^2}{2!} + \frac{2.83^3}{3!} + \frac{2.83^4}{4!} \right]
\]

\[
= .8243
\]

HW #4 due 2/6

p. 123 #48, 55 (Bin) p. 135 #50, 52
p. 130 #68, 74, 75 (Hyper, NB) (Pois)}
48. NBC News reported on May 2, 2013, that 1 in 20 children in the United States have a food allergy of some sort. Consider selecting a random sample of 25 children and let $X$ be the number in the sample who have a food allergy. Then $X \sim \text{Bin}(25, .05)$.
   a. Determine both $P(X \leq 3)$ and $P(X < 3)$.
   b. Determine $P(X \geq 4)$.
   c. Determine $P(1 \leq X \leq 3)$.
   d. What are $E(X)$ and $\sigma_X$?
   e. In a sample of 50 children, what is the probability that none has a food allergy?

55. Twenty percent of all telephones of a certain type are submitted for service while under warranty. Of these, 60% can be repaired, whereas the other 40% must be replaced with new units. If a company purchases ten of these telephones, what is the probability that exactly two will end up being replaced under warranty?

68. Eighteen individuals are scheduled to take a driving test at a particular DMV office on a certain day, eight of whom will be taking the test for the first time. Suppose that six of these individuals are randomly assigned to a particular examiner, and let $X$ be the number among the six who are taking the test for the first time.
   a. What kind of a distribution does $X$ have (name and values of all parameters)?
   b. Compute $P(X = 2)$, $P(X \leq 2)$, and $P(X \geq 2)$.
   c. Calculate the mean value and standard deviation of $X$. 
74. A second-stage smog alert has been called in a certain area of Los Angeles County in which there are 50 industrial firms. An inspector will visit 10 randomly selected firms to check for violations of regulations.

a. If 15 of the firms are actually violating at least one regulation, what is the pmf of the number of firms visited by the inspector that are in violation of at least one regulation?

b. If there are 500 firms in the area, of which 150 are in violation, approximate the pmf of part (a) by a simpler pmf.

c. For \(X\) = the number among the 10 visited that are in violation, compute \(E(X)\) and \(V(X)\) both for the exact pmf and the approximating pmf in part (b).

75. The probability that a randomly selected box of a certain type of cereal has a particular prize is .2. Suppose you purchase box after box until you have obtained two of these prizes.

a. What is the probability that you purchase \(x\) boxes that do not have the desired prize?

b. What is the probability that you purchase four boxes?

c. What is the probability that you purchase at most four boxes?

d. How many boxes without the desired prize do you expect to purchase? How many boxes do you expect to purchase?
80. Let $X$ be the number of material anomalies occurring in a particular region of an aircraft gas-turbine disk. The article "Methodology for Probabilistic Life Prediction of Multiple-Anomaly Materials" (Amer. Inst. of Aeronautics and Astronautics J., 2006: 787–793) proposes a Poisson distribution for $X$. Suppose that $\mu = 4$.
   a. Compute both $P(X \leq 4)$ and $P(X < 4)$.
   b. Compute $P(4 \leq X \leq 8)$.
   c. Compute $P(8 \leq X)$.
   d. What is the probability that the number of anomalies exceeds its mean value by no more than one standard deviation?

82. Consider writing onto a computer disk and then sending it through a certifier that counts the number of missing pulses. Suppose this number $X$ has a Poisson distribution with parameter $\mu = .2$. (Suggested in "Average Sample Number for Semi-Curtailed Sampling Using the Poisson Distribution," J. Quality Technology, 1983: 126–129.)
   a. What is the probability that a disk has exactly one missing pulse?
   b. What is the probability that a disk has at least two missing pulses?
   c. If two disks are independently selected, what is the probability that neither contains a missing pulse?