Homework # 7 *Statistics 134, Bandyopadhyay, Spring 2014*

3.2.20 Write $p(x) = P(X = x)$. Solve the system of equations

$$\begin{aligned} p(0) + p(1) + p(2) &= 1 \\ p(1) + 2p(2) &= \mu_1 \\ p(1) + 4p(2) &= \mu_2 \end{aligned}$$

for $p(2) = \frac{\mu_2 - \mu_1}{2}$, $p(1) = 2\mu_1 - \mu_2$, $p(0) = 1 - \left(\frac{\mu_2 - \mu_1}{2}\right) - (2\mu_1 - \mu_2)$.

3.3.8 a) $N = X_1 + X_2 + X_3$ b) $E(N) = \frac{1}{5} + \frac{1}{4} + \frac{1}{3} = \frac{47}{60}$

c) Note that N is the indicator of the event $A_1 \cup A_2 \cup A_3$, since A_i's are disjoint. So $Var(N) = (\frac{1}{5} + \frac{1}{4} + \frac{1}{3})(1 - \frac{1}{5} - \frac{1}{4} - \frac{1}{3}) = \frac{611}{3600}$

d) $Var(N) = Var(X_1) + Var(X_2) + Var(X_3) = \frac{1}{5} \cdot \frac{4}{5} + \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{2}{3} = \frac{2051}{3600}$

e) $Var(N) = \left[\frac{1}{5} \cdot 3^2 + (\frac{1}{4} - \frac{1}{5}) \cdot 2^2 + (\frac{1}{3} - \frac{1}{4}) \cdot 1^2\right] - (\frac{1}{5} + \frac{1}{4} + \frac{1}{3})^2 = \frac{5291}{3600}$

3.3.16 a)

x	-2	-1	0	3
$P(X = x)$	.25	.25	.25	.25

Thus $E(X) = \frac{-2-1+0+3}{4} = 0$ and $Var(X) = \frac{(-2)^2+(-1)^2+0^2+3^2}{4} - 0^2 = \frac{7}{2} = 3.5$

b) Let $S_n = \sum_{i=1}^{n} X_i$ where each X_i is one play of this game. Then we wish to know $P(S_{100} \geq 25)$. $E(S_{100}) = 100E(X) = 0$ and $Var(S_{100}) = 100Var(X) = 350$. Furthermore, the sum of 100 draws should look pretty close to normal, so

$$P(S_{100} \geq 25) \approx 1 - \Phi\left(\frac{25 - .5 - 0}{\sqrt{350}}\right) = 1 - \Phi\left(\frac{24.5}{18.71}\right) = .0952$$

3.3.28 $Var(S) = \sum_i p_i(1 - p_i) = np(1 - p) - \sum_i (p_i - p)^2$.

3.6.6 a) Consider $M = \sum_{i=1}^{n} I_i$ where I_i is the indicator of the ith ball being in the ith box. Thus $E(M) = nE(I_1) = 1$, because the chance of a match at the ith box is $1/n$ for all i.

b) Follow the argument we used in class for finding the variance of the hypergeometric distribution. You are going to need two general facts about indicators: the square of an indicator is the indicator itself, and the product of two indicators is the indicator of an intersection. You will also need to note that the chance of matches at both Box i and Box j is $1/n(n-1)$ for all $i \neq j$.

$$
\begin{aligned}
E(M^2) &= E\left(\left(\sum_{i=1}^{n} I_i\right)^2\right) \\
&= E\left(\sum_{i=1}^{n} I_i^2\right) + E\left(\sum_{i\neq j} I_i I_j\right) \\
&= E\left(\sum_{i=1}^{n} I_i\right) + E\left(\sum_{i\neq j} I_i I_j\right) \\
&= \left(\sum_{i=1}^{n} E(I_i)\right) + \left(\sum_{i\neq j} E(I_i I_j)\right) \\
&= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n} \cdot \frac{1}{n-1} \\
&= 1 + 1 = 2
\end{aligned}
$$

Thus $SD(M) = \sqrt{2-1} = 1$

c) For large n, the distribution of M is approximately Poisson(1). Intuitively, the distribution is very much like a binomial($n, \frac{1}{n}$) except for the dependence between the draws, but as the number of draws gets large the dependence between draws becomes small, and the Poisson(1) becomes a good approximation.