introduction to statistics Fall 2010

Exercise 8

1. If X is a binomial random variable with expectated value 6 and variance 2.4, find $P(X = 5)$.

   In a binomial, $E(X) = np$, $Var(X) = np(1-p)$ and

   $$P(X = i) = \frac{\mathbf{n!}}{\mathbf{i!(n-i)!}}\mathbf{p^i(1-p)^{n-i}} \tag{1}$$

   Then substituting $n = \frac{E(X)}{p}$ into Var(X) gives you $\mathbf{p = 0.6}$ and **n=10**.

   $$P(X = 5) = \frac{10!}{5!(10-5)!}0.6^5(0.4)^5 \approx \mathbf{0.200658}$$

2. Teams A and B play a series of games with the first team to win 3 games being declared the winer of the series. Suppose that team A independently wins each game with probability p. Find the conditional probability tha team A wins

   (a) the series given that it wins the first game.

   The series is a Best-of-5. Let p = P(win a game).

   $$P(win\,series|win\,1st\,game) = \frac{P(win\,series \cap win\,1st\,game)}{P(win\,1st\,game)}$$

   $$= P(Team\,A\,wins\,2\,more\,games\,before\,Team\,B\,wins\,3!)$$

   $$= \frac{\mathbf{p[p^2 + 2p^2(1-p) + 3(1-p)^2p^2]}}{\mathbf{p}}$$

   (b) the first game given that it wins the series.

   $$P(win\,first|win\,series) = \frac{P(win\,series \cap win\,first)}{P(win\,series)}$$

   Where,
   $P(win\,series \cap win\,first) = p[p^2 + 2p^2(1-p) + 3(1-p)^2p^2]$
   $P(win\,series) = p \times P(win\,series|win\,first) + (1-p) \times P(win\,series|\mathbf{lose}\,first)$
   $= P(win\,series \cap win\,1st\,game) + P(win\,series \cap \mathbf{lose}\,1st\,game)$
   $= p[p^2 + 2p^2(1-p) + 3(1-p)^2p^2] + (1-p)[p^3 + 3p^3(1-p)]$

   $$\mathbf{P(win\,first|win\,series)} = \frac{\mathbf{p[p^2 + 2p^2(1-p) + 3(1-p)^2p^2]}}{\mathbf{p[p^2 + 2p^2(1-p) + 3(1-p)^2p^2] + (1-p)[p^3 + 3p^3(1-p)]}}$$

3. Each of the members of a seven judge panel will independently make a correct decision with probability 0.7. If the panels decision is made by majority rule, what is the probability the panel makes the correct decision? Given that 4 of the judges agreed, what is the probability the jury made the correct decision?.

This is a binomial distribution. Let X = number of judges with correct decision.

Using **Equation(1)** from Q1, calculate:
$P(correct\, decision) = P(X = 7) + P(X = 6) + P(X = 5) + P(X = 4)$
$P(correct\, decision) \approx 0.873964$, majority rules!
$$P(4\, agreed) = \underbrace{\binom{7}{4} 0.7^4 0.3^3}_{\text{4 agreed correctly}} + \underbrace{\binom{7}{3} 0.7^3 0.3^4}_{\text{4 agreed wrongly}} \approx 0.324135$$

Baye's Theorem:
$$P(correct\, decision|4\, agreed) = \frac{P(4\, agreed \,|\, correct\, decision)P(correct\, decision)}{P(4\, agreed)}$$
$$P(correct\, decision|4\, agreed) = \frac{P(4\, agreed \cap correct\, decision\, reached)}{P(4\, agreed)}$$
$$P(correct\, decision|4\, agreed) = \frac{\binom{7}{4} 0.7^4 0.3^3}{\binom{7}{4} 0.7^4 0.3^3 + \binom{7}{3} 0.7^3 0.3^4} = \mathbf{0.7}$$

4. An insurance company needs to asses the risks associated with providing hurricane insurance. Between 1990 and 2006, Caribean sea was hit by 22 tropical storms or hurricanes. If tropical storms and hurricanes are independent and the mean has not changed, what is the probability of having a year in Caribe with each of the following.

   (a) No hits?

   Finding a year *within* 1990-2006 requires a binomial distribution with
   $E(X) = 22/16 = 1.375$
   Var(X) = 0 (mean *has not changed*).
   n = 2006-1990 = 16.
   $p = E(X)/n$ (property of binomial distribution) ≈ 0.085938

   Using Equation(1) from Q1:
   $P(X = 0) \approx \mathbf{0.237473}$

   (b) Exactly one hit?

   Using Equation(1) from Q1:
   $P(X = 1) \approx \mathbf{0.35722437}$

   (c) More than three hits?

   Using Equation(1) from Q1:
   $P(X > 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X - 3)$
   $P(X > 3) \approx \mathbf{0.042898}$

NOTE: This can also be done with a Poisson distribution and **be correct** if you consider a storm/hurricane in a new year. Poisson only requires one parameter (mean = 1.375 = np = E(X) of the Binomial)

Since the Poisson distribution is an approximation of the Binomial distribution (i.e., its parameter is the E(X) of Binomial), solutions are similar:
$P(X = 0) \approx 0.252839596$
$P(X = 1) \approx 0.347654444$
$P(X > 3) \approx 0.050946166$

5. A foundry ships engine blocks in lots of size 20. Since no manufacturing process is perfect, defective blocks are inevitable. However, to detect the defect, the block must be destroyed. Thus we cannot test each block. Before accepting a lot, three items are selected and tested. Suppose that a given lot actually contains five defective items.

   (a) Define X and calculate its distribution.

   X = number of blocks NOT defective out of the 3 selected and tested.
   Because blocks are selected without replacement on the available blocks in a lot (testing = destroying blocks), X is DEPENDENT on available blocks after each selection. Given data in the question:
   N = 20 total
   n = 3 selected
   m = 15 working
   $p = \frac{15}{20} = \frac{3}{4}$ healthy blocks in lot ∼**Hypergeometric distribution**

$$P(X = i) = \frac{\binom{m}{X}\binom{N-m}{n-X}}{\binom{N}{n}} = \frac{\binom{n}{X}\binom{N-n}{m-X}}{\binom{N}{m}} \tag{2}$$

   (b) Find E(X) and Var(X).

   $E(X) = np = \mathbf{2.25}$ out of 3 selected are expected to be working.
   $Var(X) = \frac{N-n}{N-1}np(1-p) \approx \mathbf{0.503289947}$

   (c) Set up the calculations needed to find $P(X < 3)$.

   Using Equation(2):
   $P(X = 0) \approx 0.00877193$
   $P(X = 1) \approx 0.131578947$
   $P(X = 2) \approx 0.460526316$
   $\mathbf{P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) \approx 0.600877193}$