Stat 134 Midterm Spring 2013

Instructor: Allan Sly

Name:

SID:

There are 4 questions worth a total of 30 points. Attempt all questions and show your working (except in Question 1). Answer the questions in the space provided. Additional space is available at the final page. No notes, cheat sheets, calculators or other references are permitted. Answers can be left in numerical form without simplification except where specified.

Question 1 True or false? No explanation is required. (1 point each)

(i) If A and B are events such that $A \subset B$ (i.e. A is a subset of B) then $P(A \cap B) = P(A)P(B)$.

False since $P(A \cap B) = P(A)$.

(ii) A bag contains red and green balls. If I draw balls *without* replacement until the first green ball then the number of draws is geometric.

False, it would be geometric is it were sampled with replacement.

(iii) Two random variables X and Y are independent with X distributed as $\text{Bin}(n_1, p)$ and Y distributed as $\text{Bin}(n_2, p)$ where $n_1 \neq n_2$. Then $X + Y$ also has a binomial distribution.

True - the distribution in $Bin(n_1 + n_2, p)$, we have $n_1 + n_2$ independent trials in total.

(iv) Two random variables X and Y are independent with X distributed as $\text{Bin}(n, p_1)$ and Y distributed as $\text{Bin}(n, p_2)$ where $p_1 \neq p_2$. Then $X + Y$ also has a binomial distribution.

False - if the trials have different probabilities then it is not binomial.

Question 2 Coin 1 has probability p_1 of heads while coin 2 has probability p_2. The coins are placed in a bag and one is drawn out at random and tossed twice. Let Y be the number of heads.

(a)[3 points] What is the probability that $Y = 1$?

(b) [3 points] Given $Y = 1$ what is the probability that the coin drawn was coin 1?

Solution

(a) We partion the event $Y = 1$ into the two possible choices of the coin. The conditional probability is a binomial.

$$\begin{aligned}
\mathbb{P}[Y=1] &= \mathbb{P}[Y=1, \mathrm{Coin1}] + \mathbb{P}[Y=1, \mathrm{Coin2}] \\
&= \mathbb{P}[Y=1 \mid \mathrm{Coin1}]\mathbb{P}[\mathrm{Coin1}] + \mathbb{P}[Y=1 \mid \mathrm{Coin2}]\mathbb{P}[\mathrm{Coin2}] \\
&= \binom{2}{1} p_1(1-p_1) \cdot \frac{1}{2} + \binom{2}{1} p_2(1-p_2) \cdot \frac{1}{2} = p_1(1-p_1) + p_2(1-p_2).
\end{aligned}$$

(b) We use Bayes rule:

$$\begin{aligned}
\mathbb{P}[\mathrm{Coin1} \mid Y=1] &= \frac{\mathbb{P}[Y=1 \mid \mathrm{Coin1}]\mathbb{P}[\mathrm{Coin1}]}{P[Y=1]} \\
&= \frac{p_1(1-p_1)}{p_1(1-p_1) + p_2(1-p_2)}.
\end{aligned}$$

Question 3 A certain lottery has weekly drawings on every Saturday. The chance of winning for each ticket is $1/100$ independently.

(a) [3 points] You buy 50 tickets. Let X be the number of winning tickets you hold. What is the distribution of X and find the probability that $X \geq 2$.

(b) [3 points] If you buy 4000 tickets give the normal approximation to the chance of winning at least 50 times (express your solution in terms of Φ the c.d.f. of the standard normal).

(c) [4 points] Sam buys 10 tickets every Friday each week until he has at least one winning ticket. Independently Harry buys 1 ticket every Friday until he has one winning ticket.

Let S be the number of weeks Sam buys tickets and let H be the number of weeks Harry buys tickets. What is the distribution of H? Find the probability (do not leave sums unsimplified) that Harry will stop buying lottery tickets before Sam (i.e. $H < S$).

Solution

(a) The distribution of X is $Bin(50, 1/100)$, it is 50 independent trials each with probability $1/100$.

$$\begin{aligned}\mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X = 0] - \mathbb{P}[X = 1] \\ &= 1 - \binom{50}{0}\left(\frac{1}{100}\right)^0\left(\frac{99}{100}\right)^{50} - \binom{50}{1}\left(\frac{1}{100}\right)^1\left(\frac{99}{100}\right)^{49}.\end{aligned}$$

(b) Let Y be the number of winning tickets we hold. Then Y has distribution $Bin(4000, 1/100)$. We have $\mathbb{E}Y = 4000 \cdot \frac{1}{100} = 40$ and $\text{SD}(Y) = \sqrt{4000 \cdot \frac{1}{100}(1 - \frac{99}{100})} = \sqrt{39.6}$. The normal approximation says that $\frac{Y-40}{\sqrt{39.6}}$ is approximately a standard normal and

$$\begin{aligned}\mathbb{P}[Y \geq 50] &= \mathbb{P}[\frac{Y - 40}{\sqrt{39.6}} \geq \frac{50 - 40}{\sqrt{39.6}}] \\ &\approx 1 - \Phi(\frac{50 - 40 - 0.5}{\sqrt{39.6}}). \\ &\approx 1 - \Phi(1.51) \approx 0.065.\end{aligned}$$

(c) H has a Geom(1/100) distribution since Harry has 1/100 chance of winning each week. Sam has $1 - \left(\frac{99}{100}\right)^{10}$ probability of winning at least once each week so S has distribution Geom$(1 - \left(\frac{99}{100}\right)^{10})$. Set $p_1 = 1/100$ and $p_2 = 1 - \left(\frac{99}{100}\right)^{10}$. Since H and S are geometric,

$$\mathbb{P}[H = k] = p_1(1 - p_1)^{k-1}, \quad \mathbb{P}[H > k] = (1 - p_1)^k, \quad \mathbb{P}[S > k] = (1 - p_2)^k.$$

Summing over the possible values which H could take and using the fact that they are independent we have

$$\begin{aligned}
\mathbb{P}[S > H] &= \sum_{k=1}^{\infty} \mathbb{P}[S > H, H = k] = \sum_{k=1}^{\infty} \mathbb{P}[S > k]\mathbb{P}[H = k] \\
&= \sum_{k=1}^{\infty} (1 - p_2)^k p_1 (1 - p_1)^{k-1} \\
&= (p_1)(1 - p_2) \sum_{k=1}^{\infty} ((1 - p_2)(1 - p_1))^{k-1} \\
&= \frac{(p_1)(1 - p_2)}{1 - (1 - p_2)(1 - p_1)} \\
&= \frac{\frac{1}{100}\left(\frac{99}{100}\right)^{10}}{1 - \left(\frac{99}{100}\right)^{11}} \approx 0.106
\end{aligned}$$

Question 4 You are dealt 13 cards from a well shuffled standard deck (52 cards where each type appears 4 times in the deck).

(a) [3 points] What is the chance of getting all 13 types (Ace, King, Queen, Jack, $10, \ldots, 2$) in the hand?

Let N be the number of distinct types in the hand.

(b) [3 points] Find $E[N]$.

(c) [4 points] Find $\text{Var}[N]$ (do not leave sums unsimplified).

Solution

(a) After getting k distinct cards from the first k draws the chance that the next one is distinct is $\frac{4(13-k)}{52-k}$ so

$$\mathbb{P}[\text{All Distinct}] = \frac{52}{52}\frac{48}{51}\frac{44}{50}\cdots\frac{4}{40} = \frac{4^{13}\cdot 13!}{52!/39!} = \frac{4^{13}}{\binom{52}{13}}$$

Alternatively the number of *unordered* hands with 1 of each type is 4^{13} - a choice of 4 possible cards for each type. The total number of unordered hands is $\binom{52}{13}$ which gives $\mathbb{P}[\text{All Distinct}] = 4^{13}/\binom{52}{13}$.

(b) Let N_i be the indicator of the event A_i that type i is present in the hand for types $i \in \{\text{Ace, King, Queen, Jack}\}$. Then

$$\mathbb{E}N = \sum_{i=1}^{13} \mathbb{E}N_1 = 13\mathbb{P}[A_i] = 13[1 - \mathbb{P}[A_i^c]].$$

The probability that i is not present is the chance of not drawing an ace from 13 draws which by the hypergeometric formula is

$$\mathbb{P}[A_i^c] = \frac{\binom{4}{0}\binom{48}{13}}{\binom{52}{13}} = \frac{\binom{48}{13}}{\binom{52}{13}}$$

Hence

$$\mathbb{E}N = 13[1 - \frac{\binom{48}{13}}{\binom{52}{13}}].$$

A lot of people chose indicators to be the event that the kth card is a distinct type - this probability is hard to calculate although you can do it.

(c) Type i and j being distinct are not independent events so we can't just add up the variances. Instead we expand the second moment to calculate it

$$\mathbb{E}[N^2] = \sum_{i=1}^{13}\sum_{i=j}^{13} \mathbb{E}N_iN_j = 13\mathbb{E}N_1^2 + 13\cdot 12\mathbb{E}N_1N_2 = 13\mathbb{E}N_1 + 156\mathbb{P}[A_1A_2]$$

Now

$$\mathbb{P}[A_1A_2] = 1 - \mathbb{P}[A_1^c] - \mathbb{P}[A_2^c] + \mathbb{P}[A_1^cA_2^c] = 1 - 2\frac{\binom{48}{13}}{\binom{52}{13}} + \frac{\binom{44}{13}}{\binom{52}{13}}$$

where for the last term we calculated the probability of getting neither type 1 or 2 - also a hypergeomtric formula. Putting it all together

$$\mathrm{Var}[N] = \mathbb{E}[N^2] - (\mathbb{E}N)^2 = 13[1 - \frac{\binom{48}{13}}{\binom{52}{13}}] + 156\left(1 - 2\frac{\binom{48}{13}}{\binom{52}{13}} + \frac{\binom{44}{13}}{\binom{52}{13}}\right) - \left(13[1 - \frac{\binom{48}{13}}{\binom{52}{13}}]\right)^2.$$