CS 70 Discrete Mathematics and Probability Theory
Spring 2021 DIS 11A

1 Geometric and Poisson

Let $X \sim \text{Geo}(p)$ and $Y \sim \text{Poisson}(\lambda)$ be independent. random variables. Compute $\mathbb{P}(X > Y)$. Your final answer should not have summations.

Solution: We condition on Y so we can use the nice property of geometric random variables that $\mathbb{P}(X > k) = (1-p)^k$, this gives

$$\begin{aligned}
P(X > Y) &= \sum_{y=0}^{\infty} P(X > Y | Y = y) \cdot P(Y = y) \\
&= \sum_{y=0}^{\infty} (1-p)^y \cdot \frac{e^{-\lambda}\lambda^y}{y!} \\
&= e^{-\lambda p} e^{\lambda p} \sum_{y=0}^{\infty} \frac{e^{-\lambda}(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p} \sum_{y=0}^{\infty} \frac{e^{-\lambda(1-p)}(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p}
\end{aligned}$$

To simplify the last summation we observed that the sum could be interpreted as the sum of the probabilities for a Poisson($\lambda(1-p)$) random variable, which is equal to 1.

2 Vegas

On the planet Vegas, everyone carries a coin. Many people are honest and carry a fair coin (heads on one side and tails on the other), but a fraction p of them cheat and carry a trick coin with heads on both sides. You want to estimate p with the following experiment: you pick a random sample of n people and ask each one to flip his or her coin. Assume that each person is independently likely to carry a fair or a trick coin.

1. Given the results of your experiment, how should you estimate p?
   (*Hint:* Construct an (unbiased) estimator for p such that $E[\hat{p}] = p$.)
2. How many people do you need to ask to be 95% sure that your answer is off by at most 0.05?

Solution:

1. We want to construct an estimate $\hat{p}$ such that $\mathbb{E}[\hat{p}] = p$. Then, if we have a large enough sample, we'd expect to get a good estimate of p. Let X_i be the indicator that the ith person's coin flips to a heads. What we observe is the fraction of people whose coin is heads. In other words, we measure $X = \frac{1}{n}\sum_{i=1}^{n} X_i$. How can we use this observation to construct $\hat{p}$?

   First,
   $$\mathbb{E}[X] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mathbb{E}[X_i] = p \cdot 1 + (1-p)\cdot\frac{1}{2},$$
   where the last equality follows from total probability. Solving for p, we find that
   $$p = 2\,\mathbb{E}[X] - 1 = \mathbb{E}[2X-1].$$
   Thus, our estimator $\hat{p}$ should be $2X-1$.

2. We want to find n such that $P[|\hat{p}-p| \le 0.05] > 0.95$. Another way to state this is that we want
   $$P[|\hat{p}-p| > 0.05] \le 0.05.$$
   Notice that $\mathbb{E}[\hat{p}] = p$ by construction, so we can immediately apply Chebyshev's inequality on $\hat{p}$. What we get is:
   $$P[|\hat{p}-p| > 0.05] \le \frac{\mathrm{Var}[\hat{p}]}{0.05^2} \le 0.05$$
   So, we want n such that $\mathrm{Var}[\hat{p}] \le 0.05^3$.
   $$\mathrm{Var}[\hat{p}] = \mathrm{Var}[2X-1] = 4\,\mathrm{Var}[X] = \frac{4}{n^2}\,\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \frac{4}{n}\,\mathrm{Var}[X_1].$$
   But X_i is an indicator (Bernoulli variable), so its variance is bounded by $\frac{1}{4}$ (note that $p(1-p)$ is maximized at $p = \frac{1}{2}$ to yield a value of $\frac{1}{4}$). Therefore we have
   $$\mathrm{Var}[\hat{p}] \le \frac{4}{n}\frac{1}{4} = \frac{1}{n}.$$
   So, we choose n such that $\frac{1}{n} \le 0.05^3$, so $n \ge \frac{1}{0.05^3} = 8000$.

3 LLSE

We have two bags of balls. The fractions of red balls and blue balls in bag A are $2/3$ and $1/3$ respectively. The fractions of red balls and blue balls in bag B are $1/2$ and $1/2$ respectively. Someone gives you one of the bags (unmarked) uniformly at random. You then draw 6 balls from that same bag with replacement. Let X_i be the indicator random variable that ball i is red. Now, let us define $X = \sum_{1 \le i \le 3} X_i$ and $Y = \sum_{4 \le i \le 6} X_i$. Find $L(Y \mid X)$. *Hint*: Recall that

$$L(Y \mid X) = \mathbb{E}(Y) + \frac{\operatorname{cov}(X,Y)}{\operatorname{Var}(X)}\big(X - \mathbb{E}(X)\big).$$

Also remember that covariance is bilinear.

Solution:

Note that although the indicator random variables are not independent, we can still apply linearity of expectation. By symmetry, we also know that each indicator follows the same distribution. Therefore:

$$\begin{aligned}
\mathbb{E}[X] = \mathbb{E}[Y] &= 3 \cdot \mathbb{E}(X_1) = 3 \cdot \mathbb{P}(X_1 = 1) = 3 \cdot \left(\frac{1}{2}\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{1}{2}\right) = \frac{7}{4}.\\
\operatorname{cov}(X,Y) &= \operatorname{cov}\Big(\sum_{1\le i\le 3} X_i, \sum_{4\le j\le 6} X_j\Big) = 9 \cdot \operatorname{cov}(X_1, X_4)\\
&= 9 \cdot \big(\mathbb{E}(X_1X_4) - \mathbb{E}(X_1)\cdot\mathbb{E}(X_4)\big).\\
\mathbb{E}(X_1X_4) - \mathbb{E}(X_1)\,\mathbb{E}(X_4) &= \mathbb{P}(X_1 = 1, X_4 = 1) - \mathbb{P}(X_1 = 1)^2\\
&= \left[\frac{1}{2}\cdot\left(\frac{2}{3}\right)^2 + \frac{1}{2}\cdot\left(\frac{1}{2}\right)^2\right] - \left[\frac{1}{2}\cdot\left(\frac{2}{3}\right) + \frac{1}{2}\cdot\left(\frac{1}{2}\right)\right]^2 = \frac{1}{144}.\\
\operatorname{Var}(X) &= \operatorname{cov}\Big(\sum_{1\le i\le 3} X_i, \sum_{1\le j\le 3} X_j\Big)\\
&= 3\cdot\operatorname{Var}(X_1) + 6\cdot\operatorname{cov}(X_1,X_2) = 3\big(\mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2\big) + 6\cdot\frac{1}{144}\\
&= 3\left[\frac{7}{12} - \left(\frac{7}{12}\right)^2\right] + 6\cdot\frac{1}{144} = \frac{111}{144}.
\end{aligned}$$

So,

$$L(Y \mid X) = \frac{7}{4} + \frac{9}{111}\left(X - \frac{7}{4}\right) = \frac{3}{37}X + \frac{119}{74}.$$