CS 174: Combinatorics and Discrete Probability Fall 2012

Homework 3

Problem 1. *(Exercise 3.5 from MU)* Given any two random variables X and Y, by the linearity of expectation we have $\mathbb{E}[X-Y]=\mathbb{E}[X]-\mathbb{E}[Y]$. Prove that, when X and Y are independent, $\mathbf{Var}[X-Y]=\mathbf{Var}[X]+\mathbf{Var}[Y]$.

Solution: From the definition of variance, we write

$$\begin{aligned}\mathrm{Var}[X-Y] &= \mathbb{E}[(X-Y)^2]-\mathbb{E}[X-Y]^2 \\ &= \mathbb{E}[X^2-2XY+Y^2]-(\mathbb{E}[X]-\mathbb{E}[Y])^2 \\ &= \mathbb{E}[X^2]-2\mathbb{E}[XY]+\mathbb{E}[Y^2]-(\mathbb{E}[X]^2-2\mathbb{E}[X]\mathbb{E}[Y]+\mathbb{E}[Y]^2) \\ &= \mathbb{E}[X^2]-\mathbb{E}[X]^2+\mathbb{E}[Y^2]-\mathbb{E}[Y]^2,\end{aligned}$$

since by independence $\mathbb{E}[XY]=\mathbb{E}[X]\mathbb{E}[Y]$. Finally, we see that

$$\mathrm{Var}[X-Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]$$

Problem 2. *(Exercise 3.15 from MU)* Let the random variable X be representable as a sum of random variables $X=\sum_{i=1}^n X_i$. Show that, if $\mathbb{E}[X_iX_j]=\mathbb{E}[X_i]\mathbb{E}[X_j]$ for every pair of i and j with $1\le i<j\le n$, then $\mathbf{Var}[X]=\sum_{i=1}^n \mathbf{Var}[X_i]$.

Solution: From the definition of variance, we write

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}\left[\left(\sum_{i=1}^n X_i-\mathbb{E}\left[\sum_{i=1}^n X_i\right]\right)^2\right] \\ &= \mathbb{E}\left[\sum_{i=1}^n\sum_{j=1}^n (X_i-\mathbb{E}[X_i])(X_j-\mathbb{E}[X_j])\right] \\ &= \mathbb{E}\left[\sum_{i=1}^n (X_i-\mathbb{E}[X_i])^2+2\sum_{i<j}(X_i-\mathbb{E}[X_i])(X_j-\mathbb{E}[X_j])\right] \\ &= \sum_{i=1}^n \mathrm{Var}(X_i)+2\sum_{i<j}\mathbb{E}[X_iX_j-\mathbb{E}[X_i]X_j-X_i\mathbb{E}[X_j]+\mathbb{E}[X_i]\mathbb{E}[X_j]] \\ &= \sum_{i=1}^n \mathrm{Var}(X_i)\end{aligned}$$

Since

$$\mathbb{E}[X_iX_j-\mathbb{E}[X_i]X_j-X_i\mathbb{E}[X_j]+\mathbb{E}[X_i]\mathbb{E}[X_j]]=2\mathbb{E}[X_i]\mathbb{E}[X_j]-2\mathbb{E}[X_i]\mathbb{E}[X_j]=0$$

Problem 3. (Exercise 3.19) Let Y be a non-negative integer-valued random variable with positive expectation. Prove

$$\frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} \leq \Pr[Y \neq 0] \leq \mathbb{E}[Y]$$

Solution: First, we consider the upper bound. By Markov's inequality, we have

$$\Pr[Y \neq 0] = \Pr[Y \geq 1] \leq \mathbb{E}[Y]$$

Consider the conditional $X = Y|Y > 0$. Recall that Jensen's inequality tells us that for any random variable X,

$$\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$$

Then, we have that

$$\mathbb{E}[Y|Y \neq 0]^2 \leq \mathbb{E}[Y^2|Y \neq 0]$$

Now we compute each side of the above inequality. For the left-hand, side we have

$$\begin{aligned}
\mathbb{E}[Y|Y \neq 0]^2 &= \left(\sum_{i=0}^{\infty} i \Pr[Y = i|Y \neq 0]\right)^2 \\
&= \left(\sum_{i=0}^{\infty} i \frac{\Pr[Y = i, Y \neq 0]}{\Pr[Y \neq 0]}\right)^2 \\
&= \left(\sum_{i=1}^{\infty} i \frac{\Pr[Y = i]}{\Pr[Y \neq 0]}\right)^2 \\
&= \frac{\mathbb{E}[Y]^2}{\Pr[Y \neq 0]^2}.
\end{aligned}$$

For the right-hand side, we have

$$\begin{aligned}
\mathbb{E}[Y^2|Y \neq 0] &= \sum_{i=0}^{\infty} i^2 \Pr[Y = i|Y \neq 0] \\
&= \sum_{i=1}^{\infty} i^2 \frac{\Pr[Y = i]}{\Pr[Y \neq 0]} \\
&= \frac{\mathbb{E}[Y^2]}{\Pr[Y \neq 0]}.
\end{aligned}$$

Putting everything together, we have

$$\begin{aligned}
\frac{\mathbb{E}[Y]^2}{\Pr[Y \neq 0]^2} &\leq \frac{\mathbb{E}[Y^2]}{\Pr[Y \neq 0]} \\
\frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} &\leq \Pr[Y \neq 0],
\end{aligned}$$

which concludes the proof.

Alternatively, we can use the Cauchy-Schwartz inequality:

$$\begin{aligned}\mathbb{E}[Y\mathbb{I}[Y>0]]^2 &\le \mathbb{E}[Y^2]\mathbb{E}[\mathbb{I}[Y>0]^2] \\ &= \mathbb{E}[Y^2]\Pr[Y>0]\end{aligned}$$

Problem 4. *(Exercise 3.20 from MU)*

(a) Chebyshev's inequality uses the variance of a random variable to bound its deviation from its expectation. We can also use higher moments. Suppose that we have a random variable X and an even integer k for which $\mathbb{E}[(X-\mathbb{E}[X])^k]$ is finite. Show that

$$\Pr\left(|X-\mathbb{E}[X]| \ge t\sqrt[k]{\mathbb{E}[(X-\mathbb{E}[X])^k]}\right) \le \frac{1}{t^k}$$

Solution: Let $Y=(X-\mathbb{E}[X])^k$. By Markov's inequality we have $\Pr[Y \ge t^k\mathbb{E}[Y]] \le \frac{\mathbb{E}[Y]}{t^k\mathbb{E}[Y]} = \frac{1}{t^k}$. Now, we have

$$\Pr\left[Y \ge t^k\mathbb{E}[Y]\right] = \Pr\left[\sqrt[k]{Y} \ge t\sqrt[k]{\mathbb{E}[Y]}\right] = \Pr\left[|X-\mathbb{E}[X]| \ge t\sqrt[k]{\mathbb{E}[(X-\mathbb{E}[X])^k]}\right]$$

where the first step is true since we take the kth root of both sides of the inequality, and the second step is true since the kth root of a number, where k is even, is the absolute value. Putting this together with the Markov's inequality, we have

$$\Pr\left[|X-\mathbb{E}[X]| \ge t\sqrt[k]{\mathbb{E}[(X-\mathbb{E}[X])^k]}\right] \le \frac{1}{t^k}$$

(b) Why is it difficult to derive a similar inequality when k is odd?

Solution: Since X is any random variable, the value $(X-\mathbb{E}[X])^k$ may be negative for odd values k. (In fact it can't be non-negative unless X is almost surely constant). Therefore Markov's inequality would not apply.

Problem 5. *(Exercise 3.21 from MU)* A fixed point of a permutation $\pi : [1,n] \to [1,n]$ is a value for which $\pi(x)=x$. Find the variance in the number of fixed points of a permutation chosen uniformly at random from all permutations. (*Hint*: Let X_i be 1 if $\pi(i)=i$, so that $\sum_{i=1}^n X_i$ is the number of fixed points. You cannot use linearity to find $\mathbf{Var}[\sum_{i=1}^n X_i]$, but you can calculate it directly.)

Solution: Let X_i be an indicator random variable for the event that $\pi(i)=i$, making i a fixed point, i.e. $X_i=1$ when i is a fixed point, and $X_i=0$ otherwise. We can easily compute the $\mathbb{E}[X]$. Let $X=\sum_{i=1}^n$ be the number of fixed points.

First, we notice that $\mathrm{Var}[X]=\mathbb{E}[X^2]-\mathbb{E}[X]^2$. Next, we compute the expectation of the number of fixed points. Since the $\mathbb{E}[X_i]=\Pr[X_i=1]=1/n$, we have

$$\mathbb{E}[X]=\mathbb{E}\left[\sum_{i=1}^n X_i\right]=\sum_{i=1}^n \mathbb{E}[X_i]=\sum_{i=1}^n(1/n)=1$$

Now, we compute the first term in the variance,

$$
\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \\
&= \left(\sum_{i=1}^{n} \mathbb{E}[X_i^2]\right) + \left(\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}[X_i X_j]\right) \\
&= \left(\sum_{i=1}^{n} \mathbb{E}[X_i]\right) + \left(\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}[X_i X_j]\right) \\
&= 1 + \left(\sum_{i=1}^{n}\sum_{j\neq i} \Pr[X_i = 1]\mathbb{E}[X_i X_j | X_i = 1]\right) \\
&= 1 + \left(\sum_{i=1}^{n}\sum_{j\neq i} \frac{1}{n}\frac{1}{(n-1)}\right) \\
&= 1 + 1 \\
&= 2
\end{aligned}
$$

The third line follows since for indicator variables $X_i^2 = X_i$. The forth line is obtained by using conditional expectation, conditioning on the event $X_i = 1$. The fifth line comes from knowing that $\Pr[X_i = 1] = 1/n$, and conditioning on $X_i = 1$, there are $n-1$ choices for mapping element j, yielding $1/(n-1)$ as the conditional probability of j being a fixed point.

Putting everything together we have

$$\mathrm{Var}[X] = 2 - 1 = 1$$

Problem 6. *(Exercise 3.25 from MU)* The weak law of large numbers states that, if $X_1, X_2, X_3, \dots$ are independent and identically distributed random variables with mean μ and standard deviation σ, then for any constant $\epsilon > 0$ we have

$$\lim_{n\to\infty} \Pr\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) = 0.$$

Use Chebychev's inequality to prove the weak law of large numbers.
Solution:

$$\mathrm{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

So by Chebychev's inequality,

$$
\begin{aligned}
\Pr\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) &\leq \frac{\sigma^2}{n\epsilon^2} \\
&\to 0 \quad \text{as } n \to \infty
\end{aligned}
$$