# Problem Set 11 Solutions

**Pset 11 Not Assigned: For Review Only**

**Reading assignment:** 6.042 Fall 1997 Lecture Notes 23–26

**Problem 1.** The couple who is having children until they get a boy and a girl knows that the expected number of children is 3. What is the variance?

**Solution**:

After the first child, the waiting time until they get a success (a child of the opposite sex) is geometrically distributed. That is the probability that the first success occurs on the $n$th try is $(1-p)^{n-1}p$ ($n-1$ failures before 1 success).

The expectation of the geometric distribution is just the mean time to failure $1/p$.

The variance is given by

$$\sum_{n=1}^{\infty} n^2(1-p)^{n-1}p - \frac{1}{p^2}$$

From earlier in the course we know that

$$\sum_{n=1}^{\infty} n^2 x^{n-1} = \frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3}$$

Setting $x = (1-p)$ we obtain

$\sum_{n=1}^{\infty} n^2(1-p)^{n-1}p = \frac{2}{p^2} - \frac{1}{p}$

Therefore the variance is $\frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p}$.

And putting $p = 1/2$ we get that the variance is 2.

**Problem 2.** Let $X$ and $Y$ be random variables whose values are independent observations drawn uniformly at random from the set $\{1, \ldots, n\}$. Compute the following quantities:

**(a)** $\text{Var}[aX + bY]$.

**Solution**:

$$\text{E}[X] = \frac{1}{n}\sum_{i=1}^{n} i = \frac{(n+1)}{2}$$

$$\text{E}(X^2) = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

so

$$\text{Var}[X] = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

$$\begin{aligned}
\text{Var}[aX + bY] &= \text{Var}[aX] + \text{Var}[bY] \\
&= a^2\text{Var}[X] + b^2\text{Var}[Y] \\
&= (a^2 + b^2)\text{Var}[X] \\
&= \frac{(a^2+b^2)}{12}(n^2-1)
\end{aligned}$$

**(b)** $\text{E}[\max(X, Y)]$.

**Solution**:

Observe that $\Pr[\max(X, Y) = i] = \Pr[X = i \text{ and } Y < i] + \Pr[X < i \text{ and } Y = i] + \Pr[X = Y = i] = \frac{1}{n}\cdot\frac{i-1}{n} + \frac{i-1}{n}\cdot\frac{1}{n} + \frac{1}{n}\cdot\frac{1}{n} = \frac{2i-1}{n^2}$.

This means that

$$\text{E}[\max(X, Y)] = \sum_{i=1}^{n}\frac{i(2i-1)}{n^2} = \frac{1}{n^2}\sum_{i=1}^{n}(2i^2 - i) = \frac{1}{n^2}(\frac{2n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}) = \frac{(n+1)(4n-1)}{6n}.$$

**(c)** $\text{E}[\min(X, Y)]$. (*Hint:* There is a way to get the answer to this using the previous part.)

**Solution**:

Observe that $\text{E}[\max(X, Y)] + \text{E}[\min(X, Y)] = \text{E}[\max(X, Y) + \min(X, Y)] = \text{E}[X + Y] = \text{E}[X] + \text{E}[Y]$. Now $\text{E}[X] = \text{E}[Y] = \frac{n+1}{2}$. so $\text{E}[\min(X, Y)] = n + 1 - \frac{(n+1)(4n-1)}{6n} = \frac{(n+1)(2n+1)}{6n}$.

**(d)** $\text{E}[|X - Y|]$.

**Solution**:

We have that $|X - Y| = \max(X, Y) - \min(X, Y)$, so by linearity of expectation $\text{E}[|X - Y|] = \text{E}[\max(X, Y)] - \text{E}[\min(X, Y)] = \frac{(n+1)(4n-1)}{6n} - \frac{(n+1)(2n+1)}{6n} = \frac{(n^2-1)}{3n}$.

**(e)** $\mathrm{Var}[|X-Y|]$.

**Solution**:

$\mathrm{Var}[|X-Y|] = \mathrm{E}[|X-Y|^2] - \mathrm{E}^2[|X-Y|]$. But $\mathrm{E}[|X-Y|^2] = \mathrm{E}[(X-Y)^2] = \mathrm{Var}[X-Y] + \mathrm{E}^2[X-Y] = \mathrm{Var}[X-Y] = 2\mathrm{Var}[X] = (n^2-1)/6$ (since $\mathrm{E}[X-Y] = \mathrm{E}[X] - \mathrm{E}[Y] = 0$).

This means that

$$\mathrm{Var}[|X-Y|] = \frac{n^2-1}{6} - (\frac{n^2-1}{3n})^2 = \frac{(n^2-1)(n^2+2)}{18n^2}.$$

**Problem 3.** **(a)** Let $X$ be a random variable whose value is an observation drawn uniformly at random from the $\{i \in \mathbb{N} \mid -n \leq i \leq n\}$. Let $Y = X^2$. Show that $X$ and $Y$ are *not* independent, but that $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.

**Solution**:

Since $Y$ is determined by $X$, there are many cases indicating where $X$ and $Y$ are not independent, e.g.,

$$\Pr[Y=0] = 1/(2n+1), \text{ but } \Pr[Y=0 \mid X=0] = 1.$$

To see that $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$, note that $\mathrm{E}[X] = 0$ and $\mathrm{E}[XY] = \mathrm{E}[X^3] = 0$ since $X$ and $X^3$ are both symmetrical.

**(b)** Show that in general for any two random variables $X$ and $Y$ that if $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$ then $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$. In particular, this shows that it is possible for nonindependent variables to have variances that exhibit linearity.

**Solution**:

We have that $\mathrm{Var}[X+Y] = \mathrm{E}[(X+Y)^2] - (\mathrm{E}[X+Y])^2 = \mathrm{E}[X^2+2XY+Y^2] - (\mathrm{E}[X]+\mathrm{E}[Y])^2 = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 + \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2 + 2\mathrm{E}[XY] - 2\mathrm{E}[X]\mathrm{E}[Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.

**Problem 4.** Suppose that someone is infected with an unknown incurable contagious disease. Everyday he encounters $n$ healthy people and infects each of them with probability $p$. Encounters are mutually independent. These people in turn meet $n$ people and infect them with probability $p$ and so on.

**(a)** Write a recurrence for the expected number of people on the $k+1$st day in terms of the expected number of people on the $k$th day. Assume that each day every sick person meets $n$ *different* persons.

**Solution**:

Let $S_k$ be a random variable giving the number infected people after $k$ days.
We can write a recurrence for $S_{k+1}$ in terms of $S_k$.

Let the sick people after day $k$ be numbered $1, \ldots, S_k$.
Let $H_i$ be a random variable giving the number of healthy people infected by sick person $i$ in one day.
Then $H_i$ has a binomial distribution with parameters $n$ and $p$.
Then:

$$S_{k+1} = S_k + \sum_{i=1}^{S_k} H_i$$

It follows from linearity of expectation that

$$\mathrm{E}[S_{k+1}] = \mathrm{E}[S_k] + \mathrm{E}[\sum_{i=1}^{S_k} H_i].$$

Then by Wald's theorem,

$$\mathrm{E}[S_{k+1}] = \mathrm{E}[S_k] + \mathrm{E}[S_k] \cdot \mathrm{E}[H]$$

where $H$ is a variable with the binomial distribution with parameters $n$ and $p$.
Thus, $\mathrm{E}[H] = np$.

So

$$\mathrm{E}[S_{k+1}] = \mathrm{E}[S_k] + \mathrm{E}[S_k] \cdot np = (1 + np)\mathrm{E}[S_k]$$

Also, $\mathrm{E}[S_0] = 1$.
So we have a recurrence involving these expectations.
Writing $e_i$ for $\mathrm{E}[S_i]$, we have:

$$e_{k+1} = (1 + np)e_k$$

and

$$e_0 = 1$$

**(b)** Solve the recurrence of part (a).
**Solution**:
$e_k = (1 + np)^k$.

**(c)** Now suppose that any sick person recovers from the disease each day with probability $r$. Find the expected number of sick people on the $k$th day.

**Solution**:

Let $B_i$ be the indicator variable for whether the $i^{th}$ sick person gets better. Then $\mathrm{E}[B_i] = r$. Then

$$S_{k+1} = S_k + \sum_{i=1}^{S_k} (H_i - B_i)$$

It follows that

$$\mathrm{E}[S_{k+1}] = \mathrm{E}[S_k] + \mathrm{E}[S_k] \cdot \mathrm{E}[H] - \mathrm{E}[S_k]\mathrm{E}[B_i]$$

and so

$$\begin{aligned} e_{k+1} &= e_k(1 + np - r) \\ &= (1 + np - r)^k \end{aligned}$$

**Problem 5.** A miner, lost in some abandoned mine, finds himself in a room containing three doors. Door 1 leads him to freedom after two hours travel; door 2 returns him to the same room after four hours travel, and door 3 returns him to the room after eight hours travel. Each time he finds himself in the room, the miner picks a door with equal probability. Let $T$ denote the time until he finds the exit.

**(a)** Define a sequence of iid random variables $X_i$ and a stopping rule $N$ such that

$$T = \sum_{i=1}^{N} X_i$$

**Solution**:

We view this as an experiment where at each step we make a choice between three equiprobable outcomes.

Let $X_i$ = number of hours to return to room on $i$th step.

Let $N$ = number of steps until good door is selected.

Then $N$ is geometrically distributed with probability $1/3$.

**(b)** Use Wald's Identity to find $\mathrm{E}[T]$

**Solution**:

$\Pr(X_i = 2) = 1/3$

$\Pr(X_i = 4) = 1/3$

$\Pr(X_i = 8) = 1/3$

$\mathrm{E}[X_i] = 14/3$

$\mathrm{E}[N] = 3$

By Wald's Identity

$$\mathrm{E}[\sum_{i=1}^{N} X_i] = 14$$

**Problem 6.** In this problem, we analyze probabilistically the 1st stage of bubble sort performed on a list of $n$ integers $\{i_1, \dots i_n\}$. We assume that all the integers are distinct.

On the first step we compare integers $i_1$ and $i_2$. If $i_1 > i_2$ we swap them otherwise we leave them as they are.

On step $k$ $(2 \leq k \leq n-1)$, we compare the new $i_k$ with $i_{k+1}$. If $i_k > i_{k+1}$ we swap else we leave them as they are.

We assume that the order of the integers is chosen uniformly from all possible orderings.

**(a)** If a number $x$ is randomly (uniformly) chosen from a set of $n$ distinct numbers, what is the probability that $x$ is the largest number in the set?

What is the probability that on the $k$th step $(2 \leq k \leq n-1)$, $i_{k+1}$ is the biggest integer to appear in the list so far?

**Solution**:

$\Pr(\text{choosing largest of } n \text{ numbers}) = 1/n$.

$\Pr(i_{k+1} \text{ in } k \text{ step is highest so far}) = 1/(k+1)$.

**(b)** Let $C_k$ be a random variable that is equal to 1 if $i_{k+1} > i_k$. What is $\mathrm{E}[C_k]$?

**Solution**:

$$\mathrm{E}[C_k] = \Pr(C_k = 1) = \Pr(\ i_k \text{ is highest so far}) = \frac{1}{k+1}.$$

**(c)** Let random variable $C = \sum_{k=1}^{n} C_k$. Give an asymptotic bound for $\mathrm{E}[C]$ as well as an interpretation of what $\mathrm{E}[C]$ is.

**Solution**:

$$\mathrm{E}[C] = \sum_k \mathrm{E}[C_k] = \sum_{k=2}^{n} \frac{1}{k} \sim \ln n.$$

$C$ is the total number of swaps, so $\mathrm{E}[C]$ is the expected number of swaps.

**Problem 7.** Consider the following 3-stage experiment:

In the first stage, I throw a 6-sided die. The number of spots that I get determines how many instances of the second stage to perform.

In the second stage, I throw a 6-sided die. This stage is repeated once for each spot I got in the first stage. The total number of spots that I get in this stage determines how many instances of the third stage to perform.

In the third stage, I throw a 6-sided die. This stage is repeated once for each spot I got in all instances of the second stage.

For example, if the outcome of the first stage is 3, then stage 2 gets played 3 times. Suppose the outcomes of the 3 instances of stage 2 are $1, 3, 4$. Then stage 3 is played $1 + 3 + 4 = 8$ times, and the score is the sum of 8 instances of stage 3.

Assume that all throws are mutually independent.

My score is the total number of spots on all the instances of the third stage.

What is my expected score?

(Analyze this as carefully as you can, using Wald's theorem. Try to be careful about uses of independence.)

**Solution**:

Let $X$ be the random variable whose value is the roll in the first stage. Let $Y_1, Y_2, \ldots, Y_X$ be the random variable whose values are the rolls in the second stage. Define $Y = Y_1 + \cdots + Y_X$. Let $Z_1, Z_2, \ldots, Z_Y$ be the random variable whose values are the rolls in the third stage. Define $Z = Z_1 + \cdots + Z_Y$. We want to find $\mathrm{E}[Z]$. Now all the rolls are mutually independent, so we can use Wald's theorem to get $\mathrm{E}[Y] = \mathrm{E}[X]\mathrm{E}[Y_i] = (7/2)^2$. Since $Y$ is determined by rolls that are independent of the rolls in the third stage, $Y$ and $Z_i$ are independent. So we apply Wald's theorem again to get $\mathrm{E}[Z] = \mathrm{E}[Y]\mathrm{E}[Z_i] = (7/2)^3$.

The answer is $(7/2)^3$.

**Problem 8.** Suppose you flip a fair coin 100 times. The coin flips are all mutually independent. What is an upper bound on the probability that the number of heads is at least 70 ...

**(a)** ... according to Markov's Inequality?

**Solution**: The expected number of heads of 50. So the probability that the number of heads is at least 70 is at most $50/70 = 0.71$.

**(b)** ... according to Chebyshev's Inequality?

**Solution**: Let $X_i$ be the random variable whose value is 1 if the $i$th coin flip is heads. Then $\mathrm{Var}[X_i] = 1/2-(1/2)^2 = 1/4$. So $\mathrm{Var}[X_1+\cdots+X_100] = 100/4 = 25$. The variance of the number of heads is $100/4 = 25$, so the standard deviation is 5. So 70 is four times the standard deviation from the mean. Since the distribution is symmetric, the probability is at most $\frac{1}{2} \cdot \frac{1}{4^2} = \frac{1}{32}$

**(c)** ... according to Chernoff's Bound?

**Solution**: We apply Chernoff's bound with $c = 70/50 = 1.4$. This gives us that $\alpha = \ln(1.4) + 1/1.4 - 1 = 0.05076$ and that the probability is at most $e^{-0.05076\cdot 1.4\cdot 50} = 0.0286$.

**Problem 9.** Let $S$ be a sample space. Say random variable $X : S \to \mathbb{N}$ has distribution

$$X = \begin{cases} 2 & \text{with probability } 1/4, \\ 3 & \text{with probability } 1/4, \\ 5 & \text{with probability } 1/2; \end{cases}$$

and random variable $Y : S \to \mathbb{N}$ has distribution

$$Y = \begin{cases} 1 & \text{with probability } 2/3, \\ 4 & \text{with probability } 1/3; \end{cases}$$

where $X$ and $Y$ are independent.

Let random variable $Z : S \to \mathbb{N}$ be given by $Z = \max(X, Y)$.

**(a)** What is the expected value of $Z$?

**Solution**:

We first compute the values of $Z$ along with their associated probabilities. Note that $P(X = x \ \&\ Y = y) = P(X = x)P(Y = y)$ since $X$ and $Y$ are independent.

| $X$ | $Y$ | $Z$ | $Z^2$ | Pr |
|---|---|---|---|---|
| 2 | 1 | 2 | 4 | $\frac{1}{6}$ |
| 3 | 1 | 3 | 9 | $\frac{1}{6}$ |
| 5 | 1 | 5 | 25 | $\frac{1}{3}$ |
| 2 | 4 | 4 | 16 | $\frac{1}{12}$ |
| 3 | 4 | 4 | 16 | $\frac{1}{12}$ |
| 5 | 4 | 5 | 25 | $\frac{1}{6}$ |

Now

$$E(Z) = 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 5 \cdot \frac{1}{3} + 4 \cdot \frac{1}{12} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{6} = 4$$

**(b)** What is the variance of $Z$?

**Solution**:

Recall that $\text{Var}Z = E(Z^2) - E(Z)^2$. So we first compute $E(Z^2)$ using the data from the above table

$$E(Z^2) = 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 25 \cdot \frac{1}{3} + 16 \cdot \frac{1}{12} + 16 \cdot \frac{1}{12} + 15 \cdot \frac{1}{6} = \frac{52}{3}$$

Finally $\text{Var}(Z) = 52/3 - 16 = \frac{4}{3}$.

**Problem 10.** **(a)** Show that if random variables $X$ and $Y$ are independent then $\text{Cov}(X, Y) = 0$

**Solution**:

First we show that if $X$ and $Y$ are independent, then $E(XY) = E(X)E(Y)$.

$$\begin{aligned}
E(XY) &= \sum_{x,y} xyP(X = x \;\&\; Y = y) \\
&= \sum_x \sum_y xyP(X = x \;\&\; Y = y) \\
&= \sum_x \sum_y xyP(X = x)P(Y = y) \\
&= \sum_x xP(X = x) \sum_y yP(Y = y) \\
&= \sum_x xP(X = x)E(Y) \\
&= E(Y) \sum_x xP(X = x) \\
&= E(X)E(Y)
\end{aligned}$$

Since $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$, $\text{Cov}(X, Y) = 0$.

**(b)** Show that having $\text{Cov}(X, Y) = 0$ does not necessarily mean that $X$ and $Y$ are independent.

**Solution**:

Let $(X, Y)$ have joint probability given by the table below:

| $X$ | $Y$ | P |
|---|---|---|
| $-1$ | 1 | $\frac{1}{3}$ |
| 0 | 0 | $\frac{1}{3}$ |
| 1 | 1 | $\frac{1}{3}$ |

Note that $X$ and $Y$ are not independent. $P(X = 1 \;\&\; Y = 1) = \frac{1}{3} \neq \frac{2}{9} = P(X = 1)P(Y = 1)$. But $E(X) = 0$ and $E(XY) = 0$ (since $XY = X$). Thus $\text{Cov}(X, Y) = 0$.