Solutions to In-Class Problems Week 14, Wed.

Problem 1. Suppose you have learned that the average graduating MIT student's total number of credits is 200.

(a) Knowing only this average, use Markov's inequality to find a best possible upper bound for the fraction of MIT students graduating with at least 235 credits. [1]

Solution. Let X be a random variable with a distribution equal to that of the graduating MIT students' credit count. We are given that $\mathrm{E}[X] = 200$. By Markov's inequality:

$$\Pr\{X \geq 235\} \leq \frac{\mathrm{E}[X]}{235} = \frac{200}{235} \approx 0.85$$

■

(b) Demonstrate that this is a best possible bound by giving a distribution for which this bound holds with equality.

Solution. The bound is attained with equality at the two-point distribution which has non-zero values only at 0 and 235, that is,

$$\begin{aligned}
\Pr\{X = 235\} &= \frac{200}{235} \\
\Pr\{X = 0\} &= \frac{35}{235} \\
\Pr\{X = x\} &= 0 \text{ for all other } x.
\end{aligned}$$

■

(c) Suppose you are now told that no student can graduate with fewer than 170 units. How does this allow you to improve your previous bound? As before, show that this is the best possible bound.

[1] Ignore the fact that there are practical limits to the amount of time a student can stay at MIT and remain sane; That is, assume that there is no bound on the number of credits a student may earn.

Solution. We can now apply Markov's inequality to the nonnegative variable $Y = X - 170$, with expectation $\mathrm{E}[Y] = \mathrm{E}[X - 170] = \mathrm{E}[X] - 170 = 30$. So,

$$\Pr\{X \geq 235\} = \Pr\{X - 170 \geq 235 - 170\} = \Pr\{Y \geq 65\}$$

Therefore:

$$\begin{aligned}
\Pr\{X \geq 235\} &= \Pr\{Y \geq 65\} \\
&\leq \frac{\mathrm{E}[Y]}{65} \\
&\leq \frac{30}{65} \approx 0.46
\end{aligned}$$

As above, we achieve an optimum (equality in the bound) when our distribution consists of two spikes: one at $(x - 170) = c - 170$, that is, $x = 235$, and one at $(x - 170) = 0$, that is, $x = 170$.

$$\begin{aligned}
\Pr\{X = 235\} &= (200 - 170)/(235 - 170) = 30/65 \\
\Pr\{X = 170\} &= 35/65 \\
\Pr\{X = x\} &= 0 \text{ for all other } x
\end{aligned}$$

■

(d) Now suppose you *further* learn that the standard deviation of the total credits per graduating student is 7. What is the Chebyshev bound on the fraction of students who can graduate with at least 235 credits?

Solution. The variance of X is the square of the standard deviation, or 49. The variance of Y is the same as that of X, by the linearity of variance. That is, $\mathrm{Var}[Y] = \mathrm{Var}[X - 170] = \mathrm{Var}[X] - \mathrm{Var}[170] = 49 - 0$. (The variance of a constant is 0).

$$\begin{aligned}
\Pr\{X \geq 235\} &= \Pr\{Y \geq 65\} \\
&= \Pr\{Y - \mathrm{E}[Y] \geq 65 - \mathrm{E}[Y]\} \\
&= \Pr\{Y - 30 \geq 35\} \\
&\leq \Pr\{|Y - 30| \geq 35\} \\
&\leq \frac{\mathrm{Var}[Y]}{35^2} \\
&\leq \frac{49}{1225} = \frac{1}{25}
\end{aligned}$$

This is a much better bound than before! ■

Problem 2. (a) Show that Markov's Theorem only applies to nonnegative random variables. That is, give an example of a random variable to which Markov's Theorem gives a *wrong* answer.

Solution. Here is one possible answer: Let R be -10 with probability $1/2$ and 10 with probability $1/2$. Then we have:

$$\mathrm{E}[R] = -10 \cdot \frac{1}{2} + 10 \cdot \frac{1}{2} = 0$$

Suppose that we now tried to compute $\Pr\{R \geq 5\}$ using Markov's Theorem:

$$\Pr\{R \geq 5\} \leq \frac{\mathrm{E}[R]}{5} = \frac{0}{5} = 0.$$

This is the wrong answer! Obviously, R is at least 5 with probability $1/2$. ■

(b) Suppose R is a random variable that is always at least -10 and has expectation 0. Since R may be negative, Markov's theorem does not apply directly. Still, use Markov's theorem to show that the probability that R is ≥ 5 is at most $2/3$.

Solution. Let $T ::= R + 10$. Now T is a nonnegative random variable with expectation $\mathrm{E}[R + 10] = \mathrm{E}[R] + 10 = 10$, so Markov's Theorem applies and tells us that $\Pr\{T \geq 15\} \leq 10/15 = 2/3$. But $T \geq 15$ iff $R \geq 5$, so $\Pr\{R \geq 5\} \leq 2/3$. ■

Problem 3. There are n people at a circular table in a Chinese restaurant. On the table, there are n different appetizers arranged on a big Lazy Susan. Each person starts munching on the appetizer directly in front of him or her. Then someone spins the Lazy Susan so that everyone is faced with a random appetizer. In class, we saw that the expected number of people that end up with the appetizer that they had originally is 1.

Let X_i be the indicator variable for the ith person getting their own appetizer back. Let S_n be the total number of people who get their own appetizer back, so $S_n = \sum_{i=1}^{n} X_i$.

(a) What is $\mathrm{E}[X_i^2]$?

Solution. $X_i = 1$ with probability $1/n$ and 0 otherwise. Thus $X_i^2 = 1$ with probability $1/n$ and 0 otherwise. So $\mathrm{E}[X_i^2] = 1/n$. ■

(b) For $i \neq j$, what is $\mathrm{E}[X_i X_j]$?

Solution. The probability that X_i and X_j are both 1 is $1/n$. Thus $X_i X_j = 1$ with probability $1/n$, and is zero otherwise. So $\mathrm{E}[X_i X_j] = 1/n$. ■

(c) What is $\mathrm{E}[S_n^2]$?

Solution.

$$\begin{aligned}
\mathrm{E}\left[S_n^2\right] &= \sum_{i,j} \mathrm{E}[X_i X_j] \\
&= n^2 \cdot \frac{1}{n} \\
&= n.
\end{aligned}$$

Alternatively, we observe directly that

$$\Pr\left\{S_n^2 = n^2\right\} \Pr\{S_n = n\} = \frac{1}{n}$$

and

$$\Pr\left\{S_n^2 = 0\right\} \Pr\{S_n = 0\} = \frac{n-1}{n},$$

so

$$\mathrm{E}\left[S_n^2\right] = n^2 \frac{1}{n} + 0 \cdot \frac{n-1}{n} = n.$$

■

(d) What is $\mathrm{Var}[S_n]$?

Solution.

$$\begin{aligned}
\mathrm{Var}[S_n] &= \mathrm{E}\left[S_n^2\right] - \mathrm{E}^2[S_n] \\
&= n - 1^2 \\
&= n - 1.
\end{aligned}$$

■

(e) Discuss the accuracy of the Chebyshev Bound on the probability that S_n is distance x from its expectation as x ranges over integers between 1 and n.

Solution. The bound $\mathrm{Var}[S_n]/x^2$ is trivial (> 1) unless $x^2 > variance S_n$, that is, unless $x \geq \left\lfloor \sqrt{n-1} + 1 \right\rfloor$. In the case that x equals this minimum value, it still gives yields a near trivial bound of $(n-1)/\left\lfloor \sqrt{n-1} + 1 \right\rfloor \approx 1$, whereas actually,

$$\Pr\{|S_n - 1| \geq x\} = \frac{1}{n}$$

for all $x \leq n-1$, and
$$\Pr\{|S_n - 1| \geq x\} = 0$$
for $x > n-1$. At $x = n-1$, the Chebyshev Bound is $(n-1)/(n-1)^2 = 1/(n-1)$ which is still a bit larger than the actual value of $1/n$. Finally, at $x = n$, the Chebyshev Bound is $(n-1)/n^2 = 1/n - 1/n^2$ whereas the actual probability is zero. ■

Problem 4. For any random variable, R, with $\mathrm{E}[R] = \mu$ and $\mathrm{Var}[R] = v$, the Chebyshev Bound says that for any real number $x > 0$,
$$\Pr\{|R - \mu| \geq x\} \leq \frac{v}{x^2}.$$
Show that for any real number, μ, and real numbers $v, x > 0$, there is an R for which the Chebyshev Bound is tight, that is,
$$\Pr\{|R| \geq x\} = \frac{v}{x^2}. \tag{1}$$
Hint: Assume $\mu = 0$ and let R be three valued with values $0, -x$, and x.

Solution. From the hint, we aim to find an R with $\mathrm{E}[R] = 0$ and $\mathrm{Var}[R] = v$ that satisfies equation (1).

Using the further hint that R takes only values $0, -x, x$, we have
$$0 = \mathrm{E}[R] = x \Pr\{R = x\} - x \Pr\{R = -x\} = x\left(\Pr\{R = x\} - \Pr\{R = -x\}\right)$$
so
$$\Pr\{R = x\} = \Pr\{R = -x\}, \tag{2}$$
since $x > 0$. Also,
$$v = \mathrm{Var}[R] = \mathrm{E}\left[R^2\right] = x^2 \Pr\{R = -x\} + x^2 \Pr\{R = x\} = 2x^2 \Pr\{R = x\},$$
so
$$\Pr\{R = x\} = \frac{v}{2x^2}.$$
This implies
$$\Pr\{R = 0\} = 1 - \Pr\{R = -x\} - \Pr\{R = x\} = 1 - \frac{v}{x^2},$$
which completely determines the distribution of R. Moreover,
$$\Pr\{|R| \geq x\} = \Pr\{R = -x\} + \Pr\{R = x\} = \frac{v}{x^2}$$
which confirms (1).

Finally, given μ, x, and v, if we let $R' ::= R + \mu$, then R' will be the desired random variable for which the Chebyshev Bound is tight. ■

Problem 5. The covariance, $\text{Cov}[X, Y]$, of two random variables, X and Y, is defined to be $\text{E}[XY] - \text{E}[X]\,\text{E}[Y]$. Note that if two random variables are independent, then their covariance is zero.

(a) Give an example to 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	$1/3$
0	0	$1/3$
1	1	$1/3$

Note that X and Y are not independent:

$$\Pr\{X = 1 \,\&\, Y = 1\} = 1/3 \neq 2/9 = \Pr\{X = 1\}\Pr\{Y = 1\}.$$

But since $XY = X$ and $\text{E}[X] = 0$, we have

$$\text{E}[X]\,\text{E}[Y] = 0 \cdot \text{E}[Y] = 0 = \text{E}[X] = \text{E}[XY].$$

Thus $\text{Cov}[X, Y] = 0$. ■

(b) Let $X_1, \ldots, X_n$ be random variables. Prove that

$$\text{Var}[X_1 + \cdots + X_n] = \sum_{i=1}^{n} \text{Var}[X_i] + 2\sum_{i<j} \text{Cov}[X_i, X_j].$$

Solution.

$$\begin{aligned}
\text{Var}[X_1 + \cdots + X_n] &= \text{E}\left[(X_1 + \cdots + X_n)^2\right] - \text{E}^2[X_1 + \cdots + X_n] \\
&= \text{E}\left[\sum_i X_i^2 + (\sum_{i<j} 2X_iX_j)\right] - \left(\sum_i \text{E}[X_i]^2 + \sum_{i<j} 2\,\text{E}[X_i]\,\text{E}[X_j]\right) \\
&= \sum_i \text{E}\left[X_i^2\right] + \sum_{i<j} 2\,\text{E}[X_iX_j] - \sum_i \text{E}[X_i]^2 - \sum_{i<j} 2\,\text{E}[X_i]\,\text{E}[X_j] \\
&= \sum_i \text{E}\left[X_i^2\right] - \text{E}[X_i]^2 + \sum_{i<j} 2(\text{E}[X_iX_j] - \text{E}[X_i]\,\text{E}[X_j]) \\
&= \sum_i \text{Var}[X_i] + 2\sum_{i<j} \text{Cov}[X_i, X_j].
\end{aligned}$$

■