Nathan Bickel
MATH 574: Section H01
Professor: Dr. Luo
October 3, 2022

MATH 574 Homework 6

Collaboration: I discussed some of the problems with Jackson Ginn, Sam Maloney, and Jack Hyatt.

Problem 1 In this problem, we will prove a one-sided Chebyshev-type bound in several steps. The goal is to fill in the details of the proof of the theorem.
(a) Let X be a random variable and let $c \in \mathbb{R}$. Prove that $V(X) = V(X + c)$.
(b) Prove that if Y is a random variable with $E(Y) = 0$, then for any constant $c \in \mathbb{R}$, $E((Y-c)^2) = V(Y)+c^2$.
(c) Prove that for any random variable X, $E(X - E(X)) = 0$.
(d) Prove that for any $c \geq 0$ and random variable X, $p(X \geq c) \leq p(X^2 \geq c^2)$.
(e) Now prove the following theorem. A brief outline is given below, but you should write a full proof. You may cite the results proven above.

Theorem 1. Let X be a random variable with variance σ^2. Then for any $k > 0$,

$$p(X - E(X) \geq k) \leq \frac{\sigma^2}{\sigma^2 + k^2}.$$

Proof outline. Set $Y = X - E(X)$. Then argue that

$$p(X - E(X) \geq k) = p(Y \geq k) \leq p((Y + x)^2 \geq (k + x)^2),$$

for any $x \geq 0$.

Use Markov's inequality (see Homework 5, #1) to obtain that

$$p((Y + x)^2 \geq (k + x)^2) \leq \frac{E((Y + x)^2)}{(k + x)^2}.$$

Conclude that

$$\frac{E((Y + x)^2)}{(k + x)^2} = \frac{\sigma^2 + x^2}{(k + x)^2}.$$

Therefore $p(X - E(X) \geq k) \leq \frac{\sigma^2+x^2}{(k+x)^2}$, but this holds for *any* x. Hence this theorem is most useful when we minimize the function $\frac{\sigma^2+x^2}{(k+x)^2}$.

Optimize (i.e., find a minimum of) the function $\frac{\sigma^2+x^2}{(k+x)^2}$ to get a bound matching the theorem. □

Solution.

(a) We have that

$$\begin{aligned}
V(X + c) &= E((X + c)^2) - [E(X + c)]^2 && \text{(definition)}\\
&= E(X^2 + 2Xc + c^2) - [E(X + c)]^2\\
&= E(X^2) + E(2Xc) + E(c^2) - [E(X) + E(c)]^2 && \text{(linearity of expectation)}
\end{aligned}$$

$$
\begin{aligned}
&= E(X^2) + E(2Xc) + E(c^2) - (E(X)^2 + 2E(X)E(c) + E(c)^2) \\
&= E(X^2) + 2cE(X) + c^2 - (E(X)^2 + 2cE(X) + c^2) && \text{(linearity of expectation)} \\
&= E(X^2) - E(X)^2 && \text{(combining like terms)} \\
&= V(X). && \text{(definition)}
\end{aligned}
$$

(b) We have that

$$
\begin{aligned}
E((Y-c)^2) &= E(Y^2 - 2Yc + c^2) \\
&= E(Y^2) - 2cE(Y) + c^2 && \text{(linearity of expectation)} \\
&= E(Y^2) - 2c(0) + c^2 && (E(Y) = 0) \\
&= E(Y^2) - [E(Y)]^2 + c^2 && (0 = 0^2) \\
&= V(Y) + c^2. && (V(Y) = E(X^2) - [E(X)]^2)
\end{aligned}
$$

It also follows that

$$
\begin{aligned}
E((Y+c)^2) &= E((Y-(-c)^2) \\
&= V(Y) + (-c)^2 \\
&= V(Y) + c^2.
\end{aligned}
$$

(c) We have that

$$
\begin{aligned}
E(X - E(X)) &= E(X) - E(E(X)) && \text{(linearity of expectation)} \\
&= E(X) - E(X) = 0.
\end{aligned}
$$

(d) Because the event $X^2 \geq c^2 = X \leq -c \cup X \geq c$, and the events $X \leq -c$ and $X \geq c$ are disjoint, we have

$$P(X \leq -c) + P(X \geq c) = P(X^2 \geq c^2).$$

Since $P(X \leq -c)$ is nonnegative, we can conclude that

$$P(X \geq c) \leq P(X^2 \geq c^2).$$

(e) Let X be a random variable with variance σ^2. Then for any $k > 0$, we claim that

$$P(X - E(X) \geq k) \leq \frac{\sigma^2}{\sigma^2 + k^2}.$$

Let $Y = X - E(X)$. Then, for any $x > 0$ we have

$$
\begin{aligned}
P(X - E(X) \geq k) &= P(Y \geq k) && \text{(by definition)} \\
&= P(Y + x \geq k + x) && \text{(algebra)} \\
&\leq P((Y+x)^2 \geq (k+x)^2). && \text{(as shown in (d))}
\end{aligned}
$$

We have from Markov's inequality that for a random variable W on a sample space S where $W(s) > 0$ for every $s \in S$, $P(W \geq a) \leq \frac{E(W)}{a}$ for all $a \in \mathbb{R}^+$. Thus, if we choose $W = (Y+x)^2$ and $a = (k+x)^2$, Markov's inequality applies because every $w \in (Y+x)^2(S)$ is positive and a is non-negative (as $x > 0$ and both results are from squares). So we continue the inequality chain with

$$P((Y+x)^2 \geq (k+x)^2) \leq \frac{E((Y+x)^2)}{(k+x)^2}.$$

We now re-introduce the variable X. We have

$$\begin{aligned}
&Y = X - E(X) \\
&\implies E(Y) = E(X - E(X)) \\
&\implies E(Y) = 0 && \text{(as shown in part (c))} \\
&\implies E((Y+x)^2) = V(Y) + x^2. && \text{(as shown in part (b))}
\end{aligned}$$

Since $V(Y) = V(X - E(X))$, $V(Y) = V(X) = \sigma^2$ as shown in part (a). So

$$\frac{E(Y+x)^2}{(k+x)^2} = \frac{\sigma^2 + x^2}{(k+x)^2}.$$

Therefore, from our chain of inequalities, we have that $p(X - E(X) \geq k) \leq \frac{\sigma^2+x^2}{(k+x)^2}$, but this holds for any x. Hence this theorem is most useful when we minimize the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{\sigma^2+x^2}{(k+x)^2}$. We use the first derivative test and solve for x:

$$\begin{aligned}
&\frac{d}{dx}\left[\frac{\sigma^2 + x^2}{(k+x)^2}\right] = 0 \\
&\implies \frac{(k+x)^2(2x) - 2(\sigma^2 + x^2)(k+x)}{(k+x)^4} = 0 \\
&\implies \frac{(k+x)(2x) - 2(\sigma^2 + x^2)}{(k+x)^3} = 0 && \text{(provided } x \neq k\text{)} \\
&\implies \frac{2(xk - \sigma^2)}{(x+k)^3} = 0 \\
&\implies xk - \sigma^2 = 0 \\
&\implies x = \frac{\sigma^2}{k}.
\end{aligned}$$

Since $k > 0$, $\frac{\sigma^2}{k}$ will always be positive. Thus, $x = 0 < \frac{\sigma^2}{k}$, so if $x = \frac{\sigma^2}{k}$ is our only minimum, $f'(x)$ will be negative for x-values less than $\frac{\sigma^2}{k}$ provided they are greater than the discontinuity at $x = -k$. Since $k > 0$, $x = 0$ satisfies this since $-k < 0 < \frac{\sigma^2}{k}$. We have $f'(x) = \frac{2(xk-\sigma^2)}{(x+k)^3}$, so

$$f'(0) = -\frac{2\sigma^2}{k^3} < 0.$$

We repeat the same process with $x = \frac{2\sigma^2}{k}$, a value guaranteed to be greater than $x = \frac{\sigma^2}{k}$. After some simplification, we obtain

$$f'\left(\frac{2\sigma^2}{k}\right) = \frac{2k^3\sigma^2}{(k^2 + 2\sigma^2)^3} > 0.$$

Thus, $x = \frac{\sigma^2}{k}$ is an absolute minimum, because $f'(x)$ is negative for values less than $\frac{\sigma^2}{k}$ and positive for values that are greater. To conclude, we evaluate

$$\begin{aligned}
f\left(\frac{\sigma^2}{k}\right) &= \frac{\sigma^2 + \left(\frac{\sigma^2}{k}\right)^2}{\left(k + \frac{\sigma^2}{k}\right)^2} \\
&= \frac{(\sigma^2k^2 + \sigma^4)/k^2}{(k^2 + \sigma^2)^2/k^2} \\
&= \frac{\sigma^2(k^2 + \sigma^2)}{(k^2 + \sigma^2)^2}
\end{aligned}$$

$$= \frac{\sigma^2}{k^2 + \sigma^2}.$$

Therefore, the minimum value of $f(x)$ is $\frac{\sigma^2}{k^2+\sigma^2}$, and we have for all $k > 0$ that

$$p(X - E(X) \geq k) \leq \frac{\sigma^2}{\sigma^2 + k^2}.$$

$\square$

Problem 2 A biased coin has probability for heads $p = 0.75$. Suppose we flip the coin $1,000$ times. Give an upper bound for the probability that we flip at least 800 heads
(a) using Markov's inequality.
(b) using Chebyshev's inequality.
(c) using Theorem 1 in the previous question.

Solution.

Let $X : S \to \mathbb{R}$ represent the number of heads after flipping the coin 1,000 times. Then, X follows a binomial distribution with $n = 1000$ and $p = \frac{3}{4}$, and thus $E(X) = np = 750$ and $V(X) = np(1-p) = \frac{375}{2}$.

(a) We have from Markov's inequality that for a random variable W on a sample space S where $W(s) > 0$ for every $s \in S$, $P(W \geq a) \leq \frac{E(W)}{a}$ for all $a \in \mathbb{R}^+$. So if we choose $W = X$ and $a = 800$, we have

$$P(X \geq 800) \leq \frac{E(X)}{800} = \frac{15}{16}.$$

(b) We have from Chebyshev's Inequality that for a random variable W on a sample space S,

$$P(|W - E(W)| \geq k) \leq \frac{V(W)}{k^2}$$

for all $k \in \mathbb{R}^+$. So if we choose $W = X$ and $k = 50$, we have

$$P(|X - E(X)| \geq 50) \leq \frac{V(X)}{2500}.$$

Substituting, we can write $P(X \leq 700 \cup X \geq 800) \leq \frac{3}{40}$, so we must have

$$P(X \geq 800) \leq \frac{3}{40}.$$

(c) We have from Theorem 1 that for a random variable W, $P(W - E(W) \geq k) \leq \frac{V(W)}{V(W)+k^2}$. So if we choose $W = X$ and $k = 50$, we have

$$P(X - E(X) \geq 50) \leq \frac{V(X)}{V(X) + 50^2}.$$

Substituting, it follows that

$$P(X \geq 800) \leq \frac{375/2}{375/2 + 2500} = \frac{3}{43}.$$

Problem 3 Let a_n be the number of strings of length n with digits in $0-9$ that contain 2 or more consecutive 0s.
(a) Find a recurrence relation for a_n.
(b) What are the initial conditions?
(c) Determine a_7.

Solution.

Let S_n be the set of strings of length n with digits in $0-9$ that contain 2 or more consecutive 0s (so $|S_n| = a_n$). Then, let S'_n be the number of strings of length n with digits in $0-9$ that do not contain 2 or more consecutive 0s, and let $a'_n = |S'_n|$. By these definitions, any string of length n with digits in $0-9$ will either be in S_n or S'_n. By the product rule, there are 10^n such possible strings, so $a_n + a'_n = 10^n$ for all $n > 0$.

(a) Let $n \in \mathbb{N}$ such that $n \geq 3$. Then, we consider the strings $\alpha_1\alpha_2\ldots\alpha_n \in S'_n$. There are two possibilities:

Case 1: $\alpha_n \neq 0$. Then, there are a'_{n-1} possibilities for $\alpha_1\alpha_2\ldots\alpha_{n-1}$ because adding a nonzero digit to the end places no additional restrictions on the first $n-1$ digits. Since there are 9 possibilities for α_n, by the product rule there are $9a'_{n-1}$ such possible strings.

Case 2: $\alpha_n = 0$. Then, $\alpha_{n-1} \neq 0$ because if it were, the last two digits would constitute consecutive 0s. So there are a'_{n-2} possibilities for $\alpha_1\alpha_2\ldots\alpha_{n-2}$, 9 possibilities for α_{n-1}, and 1 possibility for α_n. By the product rule, there are $9a'_{n-2}$ such possible strings.

So $a'_n = 9a'_{n-1} + 9a'_{n-2}$. Since we have $a'_n = 10^n - a^n$, we have $10^n - a^n = 9(10^{n-1} - a_{n-1}) + 9(10^{n-2} - a_{n-2})$. Therefore,

$$a_n = 10^n - 9(10^{n-1} - a_{n-1} + 10^{n-2} - a_{n-2}).$$

(b) For $n = 1$, there are (rather obviously) no strings of length 1 with 2 or more consecutive 0s. For $n = 2$, there is one string of length 2 with 2 or more consecutive 0s, namely 00. So $a_1 = 0$ and $a_2 = 1$.

(c) We calculate each value up to a_7:

- $a_3 = 10^3 - 9(10^{3-1} - a_{3-1} + 10^{3-2} - a_{3-2}) = 10^3 - 9(10^2 - 1 + 10^1 - 0) = 19$
- $a_4 = 10^4 - 9(10^{4-1} - a_{4-1} + 10^{4-2} - a_{4-2}) = 10^4 - 9(10^3 - 19 + 10^2 - 1) = 280$
- $a_5 = 10^5 - 9(10^{5-1} - a_{5-1} + 10^{5-2} - a_{5-2}) = 10^5 - 9(10^4 - 280 + 10^3 - 19) = 3691$
- $a_6 = 10^6 - 9(10^{6-1} - a_{6-1} + 10^{6-2} - a_{6-2}) = 10^6 - 9(10^5 - 3961 + 10^4 - 280) = 45739$
- $a_7 = 10^7 - 9(10^{7-1} - a_{7-1} + 10^{7-2} - a_{7-2}) = 10^7 - 9(10^6 - 45739 + 10^5 - 3691) = 544870$

Problem 4 Let b_n be the number of strings of length n with digits in $0-9$ that contain no two repeated numbers in a row.
(a) Determine b_1.
(b) Prove that b_n has the recurrence relation $b_n = 9b_{n-1}$ for all $n \geq 2$.
(c) Solve the recurrence relation with the initial condition in part (a).

Solution.

Let S_n be the set of strings of length n with digits in $0-9$ that contain no two repeated numbers in a row (so $|S_n| = b_n$).

(a) There are clearly no strings of length 1 that have 2 repeated numbers in a row, so any string of length 1 will work. There are 10 choices for the digit, so $b_1 = 10$.

(b) We consider the strings $\beta_1\beta_2\ldots\beta_n \in S_n$. There are b_{n-1} choices for $\beta_1\beta_2\ldots\beta_{n-1}$, since it is a string of length $n-1$ with no two repeated numbers in a row, and 9 choices for β_n since it can equal any digit other than β_{n-1} (if it did, they would constitute two repeated numbers in a row). So by the product rule, $b_n = 9b_{n-1}$.

(c) We claim that $b_n = 10(9)^{n-1}$ for $n \in \mathbb{N}$.

First, let $n = 1$. Then,

$$b_1 = 10 = 10(1) = 10(9)^0 = 10(9)^{1-1},$$

so the claim holds for $n = 1$.

Next, let $n \in \mathbb{N}$, $n \geq 1$. Assume that $b_n = 10(9)^{n-1}$. Then, we have

$$\begin{aligned}
b_{n+1} &= 9b_n && \text{(recurrence relation)} \\
&= 9(10(9)^{n-1}) && \text{(inductive hypothesis)} \\
&= 10(9)^n && \text{(rewriting product)} \\
&= 10(9)^{(n+1)-1}.
\end{aligned}$$

So if the claim holds for n, it also holds for $n + 1$. Therefore, we have that $b_n = 10(9)^{n-1}$ for all $n \in \mathbb{N}$.