Staff Solutions to Problem Set 12

Reading: Chapter 18 on Deviation from the mean

Problem 1.
If $R$ is a nonnegative random variable, then Markov’s Theorem gives an upper bound on $\Pr[R \geq x]$ for any real number $x > \text{Ex}[R]$. If $b$ is a lower bound on $R$, then Markov’s Theorem can also be applied to $R - b$ to obtain a possibly different bound on $\Pr[R \geq x]$.

(a) Show that if $b > 0$, applying Markov’s Theorem to $R - b$ gives a smaller upper bound on $\Pr[R \geq x]$ than simply applying Markov’s Theorem directly to $R$.

Solution. Define

$$T ::= R - b.$$ 

Then $T$ is a nonnegative random variable and Markov’s Theorem can therefore be applied to $T$ to give

$$\Pr[T \geq x - b] \leq \frac{\text{Ex}[T]}{x - b} = \frac{\text{Ex}[R] - b}{x - b}.$$ 

But the event $[T \geq x - b]$ is the same as $[R \geq x]$, so

$$\Pr[R \geq x] \leq \frac{\text{Ex}[R] - b}{x - b}.$$ 

So we want to show that

$$\frac{\text{Ex}[R] - b}{x - b} < \frac{\text{Ex}[R]}{x}.$$ 

Since $x$, $b$, and $x - b$ are all positive,

$$\frac{\text{Ex}[R] - b}{x - b} < \frac{\text{Ex}[R]}{x} \iff x \text{Ex}[R] - bx < x \text{Ex}[R] - b \text{Ex}[R] \iff -bx < -b \text{Ex}[R] \iff x > \text{Ex}[R].$$

But (2) is given, which shows that (1) holds, as required.

(b) What value of $b \geq 0$ in part (a) gives the best bound?

Solution. With $b \geq 0$, $R - b$ is nonnegative iff $b \leq \text{glb}(\text{range}(R))$. So for any such $b$, applying Markov’s Theorem to $R - b$ gives

$$\Pr[R \geq x] \leq \frac{\text{Ex}[R] - b}{x - b}.$$ 

It is easy to check, by simple algebra or taking derivatives with respect to $b$, that the right hand bound is strictly decreasing in $b$, so the upper bound is least when $b = \text{glb}(\text{range}(R))$. 

---

1“glb” = “greatest lower bound.”
2Differentiating this upper bound with respect to $b$ gives

$$\frac{d}{db} \left( \frac{\text{Ex}[R] - b}{x - b} \right) = \frac{\text{Ex}[R] - x}{(x - b)^2}.$$
Problem 2.
A man has a set of \( n \) keys, one of which fits the door to his apartment. He tries the keys until he finds the correct one. Give the expectation and variance for the number of trials until success when:
(a) he tries the keys at random (possibly repeating a key tried earlier).

Solution. This is a mean time to failure problem if finding a key is taken to be a “failure”. The probability, \( p \), of failure on the \( i \)th try, given “success” on the previous tries, is \( 1/n \), so if \( T \) is the number of tries to find the right key, then \( \text{Ex}[T] = 1/(1/n) = n \).
By Lemma 18.4.3,
\[
\text{Var}[T] = \frac{1-p}{p^2} = n(n-1).
\]

(b) he chooses keys randomly from among those he has not yet tried.

Solution. \( T = k \) means that the man picks the wrong key on the first trial, and he picks the wrong key on the second trial, etc, and he picks the right key on the \( k \)-th trial. Let \( K_i \) be the indicator random variable for the \( i \)th trial, namely, \( K_i = 1 \) if he picks the right key on the \( i \)th trial, and 0 otherwise. Then
\[
\text{Pr}[T = k] = \text{Pr}[K_1 = 0 & K_2 = 0 & \cdots & K_{k-1} = 0 & K_k = 1]
\]
By the Multiplication Theorem we can compute
\[
\text{Pr}[T = k] = \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \frac{n-3}{n-2} \cdots \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} = \frac{1}{n}
\]
The expectation and variance are now easy to compute from the definitions.
\[
\text{Ex}[T] = \sum_{k=1}^{n} k \cdot \text{Pr}[T = k] = \frac{1}{n} \sum_{k=1}^{n} k = \frac{n+1}{2}
\]
Since \( x > \text{Ex}[R] \) and \( x \neq b \), this derivative is negative ---and so the bound as a function of \( b \) is strictly decreasing ---for all \( 0 \leq b \leq \text{glb}(\text{range}(R)) \).
Alternatively, to prove that the bound is strictly decreasing in \( b \), suppose \( 0 < b_1, b_2 \leq \text{glb}(\text{range}(R)) \). Since \( x - b_1 > 0 \), \( x - b_2 > 0 \), and \( x > \text{Ex}[R] \),
\[
\frac{\text{Ex}[R] - b_1}{x - b_1} < \frac{\text{Ex}[R] - b_2}{x - b_2} \quad \text{IFF} \quad x \text{Ex}[R] - b_1 x - b_2 \text{Ex}[R] + b_1 b_2 < x \text{Ex}[R] - b_1 \text{Ex}[R] - b_2 x + b_1 b_2 \quad \text{IFF}
\]
\[
(b_1 - b_2) \text{Ex}[R] < (b_1 - b_2) x \quad \text{IFF} \quad b_1 - b_2 > 0
\]
\[
b_1 > b_2.
\]
Staff Solutions to Problem Set 12

Problem 3.
Peeta bakes between 1 and 2n loaves of bread to sell every day. Each day he rolls a fair, n-sided die to get a number from 1 to n, then flips a fair coin. If the coin is heads, he bakes a number of loaves of bread equal to the value on the die, and if the coin is tails, he bakes twice that many loaves.

(a) For any positive integer k ≤ 2n, what is the probability that Peeta will make k loaves of bread on any given day? (You can express your solution by cases.)

Solution. Since the dice and the coin are independent, the change of each 2n possible outcome is \( \frac{1}{2n} \). Each outcome have the form of \((c, d)\), where \( c \in \{H, T\} \) and \( 0 < d \leq n \). If \( c = H \), every \( k < n \) is possible to be the number of loaves of bread, and if \( c = T \), every even \( k \) is possible to be the number.

Therefore, there are four possible cases. If \( k \) is odd and \( k \leq n \), then the probability is \( \frac{1}{2n} \). If \( k \) is even and \( k \leq n \), then the probability is \( \frac{1}{n} \). If \( k \) is odd and \( k > n \), then the probability is 0. If \( k \) is even and \( k > n \), then the probability is \( \frac{1}{n} \).

(b) What is the expected number of loaves Peeta will bake on any given day?

Solution. Since each possible coin-dice pair outcome is equally likely, we can sum the number of loaves of bread from each outcome and then divid by the total number of outcomes.

If \( c = H \), then the sum of the \( n \) possible outcomes is \( \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \).

If \( c = T \), then the sum of the \( n \) possible outcomes is doubled, \( \sum_{i=1}^{n} 2i = n(n + 1) \).

Therefore, the expected value is \( \frac{1}{2n} \left( \frac{n(n+1)}{2} + n(n + 1) \right) = \frac{3}{4} (n + 1) \).

(c) Continuing this process, Peeta bakes bread every day for 30 days. What is the expected total number of loaves Peeta will have baked?

Solution.

\[
30 \left( \frac{3}{4} (n + 1) \right) = \frac{45}{2} (n + 1)
\]
**Problem 4.**

We have two coins: one is a fair coin, but the other produces heads with probability $3/4$. One of the two coins is picked, and this coin is tossed $n$ times.

(a) How large must $n$ be for you to be able to infer, with 95% confidence, which of the two coins had been chosen? (Get close to the minimum value of $n$ required without considering any details of the relevant distribution functions, apart from mean and variance.)

_Hint:_ Use Chebyshev’s Theorem.

**Solution.** To guess which coin was picked, set a threshold $t$ between $1/2$ and $3/4$. If the proportion of heads is less than the threshold, guess that the fair coin had been picked; otherwise, guess the biased coin.

Let the random variable $F$ be the number of heads that would appear in the first $n$ flips of the fair coin, and let $B$ denote the number of heads that would appear in the first $n$ flips of the biased coin. We must flip the coin sufficiently many times to ensure that

$$\Pr \left[ \frac{F}{n} \geq t \right] \leq 0.05$$

and

$$\Pr \left[ \frac{B}{n} < t \right] \leq 0.05$$

A natural threshold the midpoint between $1/2$ and $3/4$, namely, $t = 5/8$.

Now, $F$ has an $(n, 1/2)$-binomial distribution, so its expectation and variance are $n/2$ and $n/4$, respectively. Using Chebyshev's inequality for the fair coin,

$$\Pr \left[ \frac{F}{n} \geq \frac{5}{8} \right] = \Pr \left[ \frac{F}{n} - \frac{1}{2} \geq \frac{5}{8} - \frac{1}{2} \right] = \Pr \left[ \frac{F - n}{2} \geq \frac{n}{8} \right]$$

$$\leq \frac{\text{Var}(F)}{(n/8)^2} = \frac{n/4}{n^2/64} = \frac{16}{n}$$

The variable $B$, on the other hand, has an $(n, 3/4)$-binomial distribution, so its expectation and variance are $n(3/4)$ and $n(3/4)(1 - 3/4) = 3n/16$, respectively. Using Chebyshev’s inequality for the biased coin,

$$\Pr \left[ \frac{B}{n} < \frac{5}{8} \right] = \Pr \left[ \frac{3}{4} - \frac{B}{n} > \frac{3}{4} - \frac{5}{8} \right] = \Pr \left[ \frac{3n}{4} - B > \frac{n}{8} \right]$$

$$\leq \frac{\text{Var}(B)}{(n/8)^2} = \frac{3n/16}{n^2/64} = \frac{12}{n}$$

So, for the required confidence level, demand that

$$\frac{16}{n} \leq 0.05, \quad \frac{12}{n} \leq 0.05.$$

These hold iff $16/n \leq 0.05$, which is true iff $n \geq 320$. So knowing the results of at least 320 flips of the chosen coin will allow us to guess its identity with 95% confidence.

Because the variance of the biased coin is less than that of the fair coin, we can do slightly better if we increase our threshold a bit to about 0.634, which gives 95% confidence with 279 coin flips.
(b) Suppose you had access to a computer program that would accept any \( n \geq 0 \) and \( p \in [0, 1] \) and generate, in the form of a plot or table, the full binomial probability density and cumulative distribution functions corresponding to those parameters. How would you find the minimum number of coin flips needed to infer the identity of the chosen coin with 95% confidence? (You do not need to determine the numerical value of this minimum \( n \), but we’d be interested to know if you did.)

**Solution.** Again, we seek to determine the values of \( n \) that satisfy both (3) and (4). Using the same threshold as before, \( t = 5/8 \), it is obvious that (3) is equivalent to

\[
\text{CDF}_F((5/8)n) \geq 0.95
\]

(5)

while (4) is equivalent to

\[
\text{CDF}_B((5/8)n) \leq 0.05
\]

(6)

Knowing that \( F \) is \((n, 1/2)\)-binomially distributed and \( B \) is \((n, 3/4)\)-binomially distributed, we can use the computer program to find the smallest \( n \) that satisfies both (5) and (6).