Problem 1.
Sometimes I forget a few items when I leave the house in the morning. For example, here are probabilities that I forget various pieces of footwear:

- left sock: 0.2
- right sock: 0.1
- left shoe: 0.1
- right shoe: 0.3

(a) Let $X$ be the number of these that I forget. What is $E[X]$?

**Solution.** $X$ is the sum of the indicator variables for forgetting each item, and the expectation of an indicator is the probability of the expected event, so the expected number of events that happen is the sum of the event probabilities:

$$E[X] = 0.2 + 0.1 + 0.1 + 0.3 = 0.7$$

(b) Give a tight upper bound on the probability that I forget one or more items when no independence assumption is made about forgetting different items.

**Solution.** By the Union Bound, the probability that I forget something is at most the sum of the probabilities of forgetting each single item, namely,

$$0.1 + 0.1 + 0.2 + 0.3 = 0.7$$

In this case, Markov’s Bound yields the same answer.

(c) Use the Markov Bound to derive an upper bound on the probability that I forget 3 or more items.

**Solution.**

$$\Pr[X \geq 3] \leq \frac{E[X]}{3} = \frac{7}{30} < 0.234.$$  

In this case, a different simple argument yields a better bound. Namely, if $X \geq 3$, then at least one of right sock and left shoe must have been forgotten, so

$$\Pr[X \geq 3] \leq \Pr[\text{forgot right sock or left shoe}]$$

$$\leq \Pr[\text{forgot right sock}] + \text{forgot right left shoe}$$

$$= 0.1 + 0.1 = 0.2.$$
(d) Now suppose that I forget each item of footwear independently. Use the Chebyshev Bound to derive an upper bound on the probability that I forget two or more items.

**Solution.** Let $X_1$ be the event I bring my left sock, $X_2$ my right sock, $X_3$ my left shoe, and $X_4$ my right shoe. Then

$$\text{Ex}[X_1] = 0.2, \text{Ex}[X_2] = 0.1, \text{Ex}[X_3] = 0.1, \text{Ex}[X_4] = 0.3.$$  

Moreover, since the $X_i$ are indicator (0-1 valued) random variables, we have

$$\text{Var}[X_1] = 0.2(1 - .2) = 0.16, \text{Var}[X_2] = 0.1(1 - .1) = 0.09,$$$$\text{Var}[X_3] = 0.1(1 - .1) = 0.09, \text{Var}[X_4] = 0.3(1 - .3) = 0.21.$$

Let $X = \sum_{i=1}^{4} X_i$. Then

$$\text{Ex}[X] = \sum_{i=1}^{4} \text{Ex}[X_i] = 0.2 + 0.1 + 0.1 + 0.3 = 0.7.$$  

Since the $X_i$ are independent,

$$\text{Var}[X] = \sum_{i=1}^{4} \text{Var}[X_i] = 0.16 + 0.09 + 0.09 + 0.21 = 0.55.$$  

Now by Chebyshev’s Inequality,

$$\Pr[X \geq 2] \leq \Pr[|X - .7| \geq 1.3] = \Pr[|X - \text{Ex}[X]| \geq 1.3]$$$$\leq \frac{\text{Var}[X]}{1.3^2} = \frac{.55}{1.3^2} \leq 0.326.$$  

(e) Use Murphy’s Law to derive a lower bound on the probability that I forget one or more items.

**Solution.** Plugging into Theorem 18.7.4, the probability that I forget one or more items is at least

$$1 - e^{-\text{Ex}[X]} = 1 - e^{-0.7} = 0.503 \ldots$$

(f) I’m supposed to remember many other items, of course: clothing, watch, backpack, notebook, pencil, kleenex, ID, keys, etc. Let $X$ be the total number of items I remember. Suppose I remember items mutually independently and $\text{Ex}[X] = 36$. Use Chernoff’s Bound to give an upper bound on the probability that I remember 48 or more items.

**Solution.** By the Chernoff Bound (with $\beta(c) := c \ln c - c + 1$),

$$\Pr[X \geq 48] = \Pr[X \geq (1 + 1/3) \text{Ex}[X]] \leq e^{-\beta(4/3)\cdot36} \approx 0.1638$$

(g) Give an upper bound on the probability that I remember 108 or more items.
**Solution.** By the Chernoff Bound,
\[
\Pr[X \geq 108] = \Pr[X \geq 3 \cdot \mathbb{E}[X]] \leq e^{-\beta(3) \cdot 36} \leq e^{-46} \approx 1.05 \times 10^{-20}.
\]

**Problem 2.**
We want to store 2 billion records into a hash table that has 1 billion slots. Assuming the records are randomly and independently chosen with uniform probability of being assigned to each slot, two records are expected to be stored in each slot. Of course under a random assignment, some slots may be assigned more than two records.

(a) Show that the probability that a given slot gets assigned more than 23 records is less than \( e^{-36} \).

*Hint:* Use Chernoff’s Bound. Note that \( \beta(12) > 18 \), where \( \beta(c) := c \ln c - c + 1 \).

**Solution.** Let \( T \) be the number of records assigned to a particular slot, say the first one. So \( \mathbb{E}[T] = 2 \). Then by Chernoff
\[
\Pr[T \geq 24] = \Pr[T \geq 12 \mathbb{E}[T]] \leq e^{-\beta(12) \mathbb{E}[T]} < e^{-18 \cdot 2} = e^{-36}.
\]

(b) Show that the probability that there is a slot that gets assigned more than 23 records is less than \( e^{-15} \), which is less than 1/3,000,000. *Hint:* \( 10^9 < e^{21} \); use part (a).

**Solution.** By the Union Bound, the probability that some slot gets assigned more than 23 records is at most 1 billion times the probability that each particular slot gets assigned more than 23 records, and therefore is at most
\[
10^9 \cdot e^{-36} < e^{21} \cdot e^{-36} = e^{-15} \approx \frac{1}{3,270,000} < \frac{1}{3,000,000}.
\]

**Problem 3.**
In this problem you will check a proof of Murphy’s Law:

**Theorem** (Murphy’s Law). Let \( A_1, A_2, \ldots, A_n \) be mutually independent events, and let \( T \) be the number of these events that occur. The probability that none of the events occur is at most \( e^{-\mathbb{E}[T]} \).

To prove Murphy’s Law, note that
\[
T = T_1 + T_2 + \cdots + T_n,
\]
where \( T_i \) is the indicator variable for the event \( A_i \). Also, remember that
\[
1 + x \leq e^x
\]
for all \( x \).

(a) Justify each line in the following derivation (without looking it up in the text):
Solution. Proof.

\[
\Pr[T = 0] = \Pr[A_1 \cup A_2 \cup \cdots \cup A_n] = \Pr[A_1 \cap A_2 \cap \cdots \cap A_n] = \prod_{i=1}^{n} \Pr[A_i] = \prod_{i=1}^{n} 1 - \Pr[A_i] \leq \prod_{i=1}^{n} e^{-\Pr[A_i]} = e^{-\sum_{i=1}^{n} \Pr[A_i]} = e^{-\sum_{i=1}^{n} \text{Ex}[T_i]} = e^{-\text{Ex}[T]},
\]

\[T = 0 \text{ iff no } A_i \text{ occurs}\]

(De Morgan’s law)

(mutual independence of \(A_i\)’s)

(complement rule)

(by (2))

(exponent algebra)

(expectation of indicator variable)

((1) & linearity of expectation)

Two special cases of Murphy’s Law are worth singling out because they come up all the time.

Corollary 3.1. Suppose an event has probability \(1/m\). Then the probability that the event will occur at least once in \(m\) independent trials is approximately \(1 - 1/e \approx 63\%\). There is a \(50\%\) chance the event will occur in \(n = (\ln 2)/m \approx 0.69m\) trials.

(b) Prove Corollary 3.1.

Solution. In this case, \(\Pr[A_i] = 1/m\) for \(1 \leq i \leq n\) and

\[
\text{Ex}[\# \text{ occurrences}] = n \frac{1}{m} = \frac{n}{m}.
\]

So by Theorem 18.7.4,

\[
\Pr[\text{no occurrence}] \leq e^{-(n/m)},
\]

and therefore

\[
\Pr[\text{at least one occurrence}] \geq 1 - e^{-(n/m)}.
\]

In fact, the \(\geq\) in (3) is a tight estimate. So if the number, \(n\) of trials is \(m\), we have

\[
\Pr[\text{at least one occurrence}] \approx 1 - e^{-(m/m)} = 1 - \frac{1}{e}.
\]

If we want

\[
1 - e^{-(n/m)} \approx \Pr[\text{at least one occurrence}] \approx \frac{1}{2},
\]

then we need

\[
e^{-(n/m)} \approx \frac{1}{2},
\]

so taking log’s we conclude

\[
n \approx m \ln 2.
\]