1. Three professors, $A$, $B$, and $C$ are coming up for review. It is common knowledge that one of them will be fired the next day and the others will be retained. Only the dean knows which one will be fired. Professor $A$ asks the chair a favor: “Please ask the Dean who will be fired, and take a message to one of my friends $B$ or $C$ to let her know that she will be retained in the morning.” The chair agrees and comes back later and tells $A$ that he gave the retain message to $B$.

What are $A$’s chances of being fired given this information? Give a mathematical explanation not just an intuitive one.

2. An experiment consists of picking a random bit string of length 5. Consider the following events.
   - $E_1$: the bit string chosen begins with 1;
   - $E_2$: the bit string chosen ends with 1;
   - $E_3$: the bit string chosen has exactly three 1s;

   (a) Find $\Pr(E_1|E_3)$.

   (b) Find $\Pr(E_3|E_2)$.

   (c) Find $\Pr(E_2|E_3)$.

   (d) Find $\Pr(E_3|E_1 \cap E_2)$.

   (e) Are $E_1$ and $E_2$ independent?

   (f) Are $E_2$ and $E_3$ independent?
3. Show:

\[ \text{Var}(X) \overset{\text{def}}{=} E((X - \mu)^2) = E(X^2) - \mu^2 \]

4. For any random variable \( X \), prove the following:

(a) For any constant \( c \), we have

\[ \text{Var}(cX) = c^2 \text{Var}(X). \]

(b) For independent random variables \( X, Y \), we have

\[ \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y). \]
5. Prove Markov's Inequality: For a non-negative random variable $X$ with expectation $E[X] = \mu$, and $\forall \alpha > 0$
\[ \Pr[X \geq \alpha] \leq \frac{E(X)}{\alpha} \]

6. Now, given Markov’s Inequality, prove Chebyshev’s Inequality: For a random variable $X$ with expectation $E(X) = \mu$, and for any $\alpha > 0$,
\[ \Pr[|X - \mu| \geq \alpha] \leq \frac{\text{Var}(X)}{\alpha^2}. \]

7. What is the probability of deviating from the mean by more than 2 standard deviations?

8. Let $X$ be the number of Heads in $n$ tosses of a fair coin. The probability that $X$ deviates from $\mu = \frac{n}{2}$ by more than $\sqrt{n}$ is at most $\frac{1}{4}$. What is the probability that it deviates by more than $5\sqrt{n}$?

9. For our random walk described in class on Monday, what is the probability that we end up 10000 steps away from our starting point after $10^6$ steps?
10. We want to estimate the proportion $p$ of Duke fans in the US population (size $n$), by taking a small random sample. How large does our sample have to be to guarantee that our estimate will be within (say) 10% (in relative terms) of the true value with probability at least 0.95?
11. Consider the problem of identifying a message as being spam. A message consists of a set of
tokens and each of a set of training messages has been classified as either spam or ham (not spam).

(a) Naive Bayes models are often used for identifying spam and other text categorization.
The Naive Bayes model states the following
\[
\Pr(Cause, Effect_1, \ldots, Effect_n) = \Pr(Cause) \prod_i (Effect_i | Cause)
\]

Why is it naive?

(b) Describe how the model can be constructed. What are the causes and effects?

(c) Explain precisely how to categorize a new document.

(d) Is the independence assumption reasonable in this case? How can the model be augmented?