1. We prove that if $Z$ is a Poisson random variable of mean $\mu$, where $\mu \geq 1$ is an integer, then $\Pr(Z \geq \mu) \geq 1/2$ and $\Pr(Z \leq \mu) \geq 1/2$.

(a) Show that $\Pr(Z = \mu + h) \geq \Pr(Z = \mu - h - 1)$ for $0 \leq h \leq \mu - 1$.

**Ans.** For any non-negative integer $k$, $\Pr(Z = k) = e^{-\mu} \mu^k/k!$ by definition. So,

\[
\frac{\Pr(Z = \mu + h)}{\Pr(Z = \mu - h - 1)} = \frac{\mu^{2k+1}}{(\mu - h)(\mu - (h - 1)) \cdots (\mu + h)} = \frac{\mu^2}{\mu^2 - h^2} \cdot \frac{\mu^2}{\mu^2 - (h - 1)^2} \cdots \frac{\mu^2}{\mu^2 - 1^2} \cdot \frac{\mu}{\mu} \geq 1.
\]

In other words, $\Pr(Z = \mu + h) \geq \Pr(Z = \mu - h - 1)$.

(b) Using part (a), argue that $\Pr(Z \geq \mu) \geq 1/2$.

**Ans.** First, recall that

\[
1 = \sum_{k=0}^{\infty} \Pr(Z = k).
\]

However, the right-side can be expressed by

\[
\sum_{k=1}^{\infty} \Pr(Z = k) = \sum_{k=0}^{\mu-1} \Pr(Z = k) + \Pr(Z \geq \mu).
\]

Then by part (a), we know that

\[
\sum_{k=0}^{\mu-1} \Pr(Z = k) \leq \sum_{k=\mu}^{2\mu-1} \Pr(Z = k) \leq \sum_{k=\mu}^{\infty} \Pr(Z = k) = \Pr(Z \geq \mu).
\]

Combining, we have

\[
1 = \sum_{k=0}^{\infty} \Pr(Z = k) = \sum_{k=0}^{\mu-1} \Pr(Z = k) + \Pr(Z \geq \mu) \leq 2 \Pr(Z \geq \mu).
\]

Thus, $\Pr(Z \geq \mu) \geq 1/2$.

2. In Page 15 of Lecture Notes 14, we showed that for any nonnegative function $f$,

\[
E \left[ f(Y_1^{(m)}, \ldots, Y_n^{(m)}) \right] \geq E \left[ f(X_1^{(m)}, \ldots, X_n^{(m)}) \right] \Pr \left( \sum Y_i^{(m)} = m \right).
\]
(a) Now, suppose we further know that \( E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \) is monotonically increasing in \( m \). Show that

\[
E \left[ f(Y^{(m)}_1, \cdots, Y^{(m)}_n) \right] \geq E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \Pr \left( \sum Y^{(m)}_i \geq m \right).
\]

**Proof.**

\[
E \left[ f(Y^{(m)}_1, \cdots, Y^{(m)}_n) \right] = \sum_{k \geq 0} E \left[ f(Y^{(m)}_1, \cdots, Y^{(m)}_n) \mid \sum Y^{(m)}_i = k \right] \Pr \left( \sum Y^{(m)}_i = k \right)
\]

\[
\geq \sum_{k \geq m} E \left[ f(Y^{(m)}_1, \cdots, Y^{(m)}_n) \mid \sum Y^{(m)}_i = k \right] \Pr \left( \sum Y^{(m)}_i = k \right)
\]

\[
= \sum_{k \geq m} E \left[ f(X^{(k)}_1, \cdots, X^{(k)}_n) \mid \sum Y^{(m)}_i = k \right] \Pr \left( \sum Y^{(m)}_i = k \right)
\]

\[
\geq \sum_{k \geq m} E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \mid \sum Y^{(m)}_i = k \right] \Pr \left( \sum Y^{(m)}_i = k \right)
\]

\[
= E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \Pr \left( \sum Y^{(m)}_i \geq m, \sum Y^{(m)}_i \geq m \right),
\]

where Line 4 follows from the assumption that \( E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \) is monotonically increasing.

(b) Combining part (a) with the results in Question 1, prove the monotonically increasing case of theorem in Page 20 of Lecture Notes 14.

**Proof.** Let \( Z = \sum Y^{(m)}_i \). Since each \( Y^{(m)}_i \) is a Poisson random variable, their sum \( Z \) is also a Poisson random variable. Further, the mean value of \( Z \) is \( m \). Thus, by results in Question 1, \( \Pr(Z \geq m) \geq \frac{1}{2} \). Combining this with part (a), we have

\[
E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \leq 2 E \left[ f(Y^{(m)}_1, \cdots, Y^{(m)}_n) \right]
\]

when \( E \left[ f(X^{(m)}_1, \cdots, X^{(m)}_n) \right] \) is monotonically increasing in \( m \). This completes the proof of the monotonically increasing case of the desired theorem.

3. Bloom filters can be used to estimate set differences. Suppose you have a set \( X \) and I have a set \( Y \), both with \( n \) elements. For example, the sets might represent our 100 favorite songs. We both create Bloom filters of our sets, using the same number of bits \( m \) and the same \( k \) hash functions. Determine the expected number of bits where our Bloom filters differ as a function of \( m, n, k, \) and \( |X \cap Y| \).

**Ans.** Let \( Z \) be a random variable denoting the number of bits where the Bloom filters differ. Let \( Z_i \) be an indicator such that

\[
Z_i = 1 \quad \text{if the } i \text{th bit of the Bloom filters differ}
\]

\[
Z_i = 0 \quad \text{otherwise.}
\]
Thus, $Z = Z_1 + Z_2 + \cdots + Z_m$

When $|X \cap Y| = r$, $Z_i = 1$ only happens when each of the $r$ common elements are not mapped to the $i$th bit, together with exactly one of the following cases (that causes the $i$th bit different):

(a) Some elements of $X - (X \cap Y)$ is mapped to the $i$th bit, but all elements of $Y - (X \cap Y)$ are not;

(b) Some elements of $Y - (X \cap Y)$ is mapped to the $i$th bit, but all elements of $X - (X \cap Y)$ are not.

Let $Q_i$ denote the event that the $r$ common elements are not mapped to the $i$th bit. By assuming that the hash functions we choose will map elements independently and uniformly at random to one of the $m$ bits, we have

$$
Pr(Z_i = 1) = Pr(Q_i \cap (\text{Case (a) or Case (b)}))
= Pr(Q_i)Pr(\text{Case (a) or Case (b)})
= Pr(Q_i)(Pr(\text{Case (a)}) + Pr(\text{Case (b)}))
= \left(1 - \frac{1}{m}\right)^r(Pr(\text{Case (a)}) + Pr(\text{Case (b)}))
= \left(1 - \frac{1}{m}\right)^r \times 2 \times \left(1 - \left(1 - \frac{1}{m}\right)^{(n-r)k}\right) \left(\frac{1}{m}\right)^{(n-r)k}
\begin{cases}
\text{some elements mapped to } i\text{th bit} \\
\text{no elements mapped to } i\text{th bit}
\end{cases}
= 2 \left(1 - \frac{1}{m}\right)^n \left(1 - \left(1 - \frac{1}{m}\right)^{(n-r)k}\right).
$$

As $Z_i$ is an indicator, $E[Z_i] = Pr(Z_i = 1)$. Thus,

$$
E[Z] = \sum_{i=0}^{m-1} E[Z_i] = m \cdot E[Z_i] = 2m \left(1 - \frac{1}{m}\right)^n \left(1 - \left(1 - \frac{1}{m}\right)^{(n-r)k}\right).
$$

4. For the leader election problem briefly introduced in Lecture Notes 15, we have $n$ users, each with an identifier. Suppose that we have a good hash function (that looks uniform and independent), which outputs a $b$-bit hash value for each identifier. One way to solve the leader election problem is as follows: Each user obtains the hash value from its identifier, and the leader is the user with the smallest hash value.

Give lower and upper bounds on the number of bits $b$ necessary to ensure that a unique leader is successfully chosen with probability $p$. Make your bounds as tight as possible.

**Ans.** (Lower Bound:) First, a unique leader is determined if and only if for some bit $i$, it is mapped exactly by one user, while all other users are mapped to the bits larger than $i$. As }
This implies that:

\[
p = \sum_{i=0}^{2^b-2} \Pr(\text{“bit } i \text{ is mapped exactly by one user” } \cap \text{ “all other users are larger than } i\text{”})
\]

\[
= \sum_{i=0}^{2^b-2} \frac{1}{2^b} \cdot \left(1 - \frac{i+1}{2^b}\right)^{n-1}
\]

\[
\leq \sum_{i=0}^{2^b-2} \frac{1}{2^b} \cdot \left(1 - \frac{1}{2^b}\right)^{n-1}
\]

\[
\leq \left(1 - \frac{1}{2^b}\right)^{n-1}.
\]

By re-arranging terms, we have:

\[
b \geq \log_2 \left(\frac{1}{1 - p^{\frac{1}{n-1}}} \right).
\]

(Upper Bound:) Next, \(p\) is greater than the probability that user 1 is the unique leader. This happens when user 1 is mapped to uniquely to some number, and this number is smallest among all numbers mapped by other users. Thus,

\[
p \geq \Pr(\text{“user 1 is mapped uniquely” } \cap \text{ “this number is smallest”})
\]

\[
= \Pr(\text{“this number is smallest” } | \text{ “user 1 is mapped uniquely”})
\times \Pr(\text{“user 1 is mapped uniquely”})
\]

\[
= \Pr(\text{“this number is smallest” } | \text{ “user 1 is mapped uniquely”}) \left(1 - \frac{1}{2^b}\right)^{n-1}
\]

\[
\geq \frac{1}{n} \left(1 - \frac{1}{2^b}\right)^{n-1}.
\]

By re-arranging terms, we have:

\[
b \leq \log_2 \left(\frac{1}{1 - (np)^{\frac{1}{n-1}}} \right).
\]

In conclusion,

\[
\log_2 \left(\frac{1}{1 - p^{\frac{1}{n-1}}} \right) \leq b \leq \log_2 \left(\frac{1}{1 - (np)^{\frac{1}{n-1}}} \right).
\]