Chernoff Bounds (Fundamentals)

Pavlos Eirinakis

1Lane Department of Computer Science and Electrical Engineering
West Virginia University

14 February, 2012
1. Tail bounds
1 Tail bounds

2 Moment Generating Functions
1. Tail bounds
2. Moment Generating Functions
3. Poisson Trials
Outline

1. Tail bounds
2. Moment Generating Functions
3. Poisson Trials
4. Chernoff Bounds
The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.
The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.
The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.

Example
Consider the experiment of tossing a fair coin $n$ times.
Note

The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.

Example

Consider the experiment of tossing a fair coin $n$ times. What is the probability that the number of heads exceeds $\frac{3}{4} \cdot n$?
Markov’s inequality
Markov's inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,
Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c}$$
Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$ P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c} $$
Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c}$$

Chebyshev's inequality
Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c}$$

Chebyshev’s inequality

Let $X$ be a random variable (not necessarily positive). Then,
Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c}$$

Chebyshev's inequality

Let $X$ be a random variable (not necessarily positive). Then,

$$P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$$
### Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c}$$

### Chebyshev's inequality

Let $X$ be a random variable (not necessarily positive). Then,

$$P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2} \quad \text{or} \quad P(|X - E[X]| \geq aE[X]) \leq \frac{\text{Var}[X]}{a^2(E[X])^2}$$
### Markov’s inequality

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E(X)}{c} \quad \text{or} \quad P(X \geq cE[X]) \leq \frac{1}{c}$$

### Chebyshev’s inequality

Let $X$ be a random variable (not necessarily positive). Then,

$$P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2} \quad \text{or} \quad P(|X - E[X]| \geq aE[X]) \leq \frac{\text{Var}[X]}{a^2(E[X])^2}$$

### Example

Consider the experiment of tossing a fair coin $n$ times. What is the probability that the number of heads exceeds $\frac{3}{4} \cdot n$?
The moment-generating function of a random variable $X$
The moment-generating function of a random variable $X$

The moment-generating function of a random variable is an alternative specification of its probability distribution.
The moment-generating function of a random variable $X$

The moment-generating function of a random variable is an alternative specification of its probability distribution.

The moment-generating function of a random variable $X$ is:

$$M_X(t) = E[e^{tX}]$$
The moment-generating function of a random variable $X$

The moment-generating function of a random variable is an alternative specification of its probability distribution. The moment-generating function of a random variable $X$ is:

$$M_X(t) = E[e^{tX}]$$

Note that $M_X(0)$ always exists and that $M_X(0) = 1$. 
The moment-generating function of a random variable $X$

The moment-generating function of a random variable is an alternative specification of its probability distribution.

The moment-generating function of a random variable $X$ is:

$$M_X(t) = E[e^{tX}]$$

Note that $M_X(0)$ always exists and that $M_X(0) = 1$. 
Theorem

Let $X$ be a random variable with moment generating function $M_X(t)$. 
Theorem

Let $X$ be a random variable with moment generating function $M_X(t)$. Under the assumption that exchanging the expectation and differentiation operands is legitimate,
Theorem

Let $X$ be a random variable with moment generating function $M_X(t)$. Under the assumption that exchanging the expectation and differentiation operands is legitimate, for all $n > 1$ we have

$$E[X^n] = M_X^{(n)}(0)$$

where $M_X^{(n)}(0)$ is the $n^{th}$ derivative of $M_X(t)$ evaluated at $t = 0$. 
Theorem

Let $X$ be a random variable with moment generating function $M_X(t)$. Under the assumption that exchanging the expectation and differentiation operands is legitimate, for all $n > 1$ we have

$$E[X^n] = M_X^{(n)}(0)$$

where $M_X^{(n)}(0)$ is the $n^{th}$ derivative of $M_X(t)$ evaluated at $t = 0$.

Proof.

Exercise.
Consider a Bernoulli random variable $X$. 
Bernoulli random variables

Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:
Bernoulli random variables

Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = \quad =$$
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}]$$
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_{x} P(X) e^{tx}$$
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_{X} P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0$$
Bernoulli random variables

Consider a Bernoulli random variable \( X \).

The moment-generating function of \( X \) is:

\[
M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = \frac{e^t}{2} + \frac{1}{2}
\]
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = e^t \cdot \frac{1}{2} + \frac{1}{2}$$

Example

Using the moment generating function of $X$: 

---

**Bernoulli random variables**

Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = e^t \cdot \frac{1}{2} + \frac{1}{2}$$

---

**Example**

Using the moment generating function of $X$: 

---
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = e^t + \frac{1}{2}.$$ 

**Example**

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$. 

Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = e^t + \frac{1}{2}$$

**Example**

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.
- Evaluate the variance of $X$, i.e., $Var(X)$. 
Consider a Bernoulli random variable $X$. The moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}] = \sum_X P(X)e^{tX} = P(X = 1)e^t + P(X = 0)e^0 = \frac{e^t}{2} + \frac{1}{2}$$

**Example**

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.
- Evaluate the variance of $X$, i.e., $Var(X)$.

Hint: Recall that $E[X^n] = M_X^{(n)}(0)$. 
Geometric random variables

Consider a geometric random variable $X$ with parameter $p$. 
Consider a geometric random variable $X$ with parameter $p$.

What is a geometric random variable?
Consider a geometric random variable $X$ with parameter $p$. What is a geometric random variable? What is its probability distribution?
Consider a geometric random variable $X$ with parameter $p$. What is a geometric random variable? What is its probability distribution? Recall that the probability distribution of $X$ on $n = 1, 2, \ldots$ is

$$P(X = n) = (1 - p)^{n-1} p$$
Consider a geometric random variable $X$ with parameter $p$.

What is a geometric random variable?
What is its probability distribution?
Recall that the probability distribution of $X$ on $n = 1, 2, \ldots$ is

$$P(X = n) = (1 - p)^{n-1}p$$

Also the expectation of functions is:

$$E[g(X)] = \sum_X P(X) \cdot g(X)$$
Geometric random variables

Consider a geometric random variable $X$ with parameter $p$.
What is a geometric random variable?
What is its probability distribution?
Recall that the probability distribution of $X$ on $n = 1, 2, \ldots$ is

$$P(X = n) = (1 - p)^{n-1}p$$

Also the expectation of functions is:

$$E[g(X)] = \sum_X P(X) \cdot g(X)$$

For geometric random variable $X$ and for $t < -\ln(1 - p)$, the moment-generating function of $X$ is:
Consider a geometric random variable $X$ with parameter $p$.

What is a geometric random variable?
What is its probability distribution?

Recall that the probability distribution of $X$ on $n = 1, 2, \ldots$ is

$$P(X = n) = (1 - p)^{n-1}p$$

Also the expectation of functions is:

$$E[g(X)] = \sum_X P(X) \cdot g(X)$$

For geometric random variable $X$ and for $t < -\ln(1 - p)$, the moment-generating function of $X$ is:

$$M_X(t) = E[e^{tx}]$$
Geometric random variables

Consider a geometric random variable $X$ with parameter $p$.
What is a geometric random variable?
What is its probability distribution?
Recall that the probability distribution of $X$ on $n = 1, 2, \ldots$ is

$$P(X = n) = (1 - p)^{n-1}p$$

Also the expectation of functions is:

$$E[g(X)] = \sum_{X} P(X) \cdot g(X)$$

For geometric random variable $X$ and for $t < -\ln(1 - p)$, the moment-generating function of $X$ is:

$$M_X(t) = E[e^{tX}]$$
$$= \frac{p}{1 - p}((1 - (1 - p)e^{t})^{-1} - 1).$$
Hence, for $t < -\ln(1 - p)$:

$$M_X(t) = \frac{p}{1 - p} \left( (1 - (1 - p)e^t)^{-1} - 1 \right).$$
Geometric random variables

Hence, for $t < -\ln(1 - p)$:

$$M_X(t) = \frac{p}{1 - p} \left( \frac{1 - (1 - p)e^t}{(1 - (1 - p)e^t)^{-1}} - 1 \right).$$

Example

Using the moment generating function of $X$: 
### Geometric random variables

Hence, for $t < -\ln(1-p)$:

\[
M_X(t) = \frac{p}{1-p}((1 - (1 - p)e^t)^{-1} - 1).
\]

### Example

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$. 

---

**Eirinakis**

Chernoff bounds
Geometric random variables

Hence, for $t < -\ln(1 - p)$:

$$M_X(t) = \frac{p}{1 - p}((1 - (1 - p)e^t)^{-1} - 1).$$

Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.
- Evaluate the variance of $X$, i.e., $Var(X)$. 
### Geometric random variables

Hence, for $t < -\ln(1 - p)$:

\[
M_X(t) = \frac{p}{1 - p} \left( (1 - (1 - p)e^t)^{-1} - 1 \right).
\]

### Example

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.
- Evaluate the variance of $X$, i.e., $Var(X)$.

### The derivatives

The first and second derivatives of $M_X(t)$ are:
### Geometric random variables

Hence, for $t < -\ln(1 - p)$:

$$M_X(t) = \frac{p}{1 - p}((1 - (1 - p)e^t)^{-1} - 1).$$

### Example

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.
- Evaluate the variance of $X$, i.e., $Var(X)$.

### The derivatives

The first and second derivatives of $M_X(t)$ are:

$$M_X^{(1)}(t) = p(1 - (1 - p)e^t)^{-2}e^t$$
Moment Generating Functions

Geometric random variables

Hence, for \( t < -\ln(1 - p) \):

\[
M_X(t) = \frac{p}{1 - p} \left( \frac{1 - (1 - p)e^t}{1 - (1 - p)e^t} \right)^{-1} - 1.
\]

Example

Using the moment generating function of \( X \):
- Evaluate the expectation of \( X \), i.e., \( E(X) \).
- Evaluate the variance of \( X \), i.e., \( Var(X) \).

The derivatives

The first and second derivatives of \( M_X(t) \) are:

\[
M_X^{(1)}(t) = p(1 - (1 - p)e^t)^{-2}e^t
\]
\[
M_X^{(2)}(t) = 2p(1 - p)(1 - (1 - p)e^t)^{-3}e^{2t} + p(1 - (1 - p)e^t)^{-2}e^t
\]
Moment Generating Functions

Geometric random variables

Hence, for \( t < -\ln(1 - p) \):

\[
M_X(t) = \frac{p}{1 - p} \left( \frac{(1 - (1 - p) e^t)^{-1} - 1}{(1 - (1 - p) e^t)^{-1}} \right).
\]

Example

Using the moment generating function of \( X \):

- Evaluate the expectation of \( X \), i.e., \( E(X) \).
- Evaluate the variance of \( X \), i.e., \( \text{Var}(X) \).

The derivatives

The first and second derivatives of \( M_X(t) \) are:

\[
M_X^{(1)}(t) = p(1 - (1 - p)e^t)^{-2} e^t
\]

\[
M_X^{(2)}(t) = 2p(1 - p)(1 - (1 - p)e^t)^{-3} e^{2t} + p(1 - (1 - p)e^t)^{-2} e^t
\]
**Theorem**

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.
**Theorem**

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

**Theorem**

If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$
Theorem

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

Theorem

If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Proof.

Theorem

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

Theorem

If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Proof.

$$M_{X+Y}(t) = E[e^{t(X+Y)}]$$
Theorem
Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

Theorem
If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Proof.

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}]$$
Theorem

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

Theorem

If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Proof.

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}]$$
### Theorem

Let $X$ and $Y$ be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then $X$ and $Y$ have the same distribution.

### Theorem

If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

### Proof.

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}]$$

$$= E[e^{tX}]E[e^{tY}]$$

$$= M_X(t)M_Y(t)$$
We will develop Chernoff bounds for the tail distribution of Poisson trials.
We will develop Chernoff bounds for the tail distribution of Poisson trials.

Poisson trials:
We will develop Chernoff bounds for the tail distribution of Poisson trials. Poisson trials:

- Sum of independent 0 – 1 random variables.
We will develop Chernoff bounds for the tail distribution of Poisson trials. **Poisson trials:**

- Sum of independent 0–1 random variables.
- The distribution of the random variables in Poisson trials are not necessarily identical.
We will develop Chernoff bounds for the tail distribution of Poisson trials. Poisson trials:

- Sum of independent 0 – 1 random variables.
- The distribution of the random variables in Poisson trials are not necessarily identical.
- Bernoulli trials are a special case of Poisson trials where the independent 0 – 1 random variables have the same distribution.
We will develop Chernoff bounds for the tail distribution of Poisson trials.

Poisson trials:

- Sum of independent 0 – 1 random variables.
- The distribution of the random variables in Poisson trials are not necessarily identical.
- Bernoulli trials are a special case of Poisson trials where the independent 0 – 1 random variables have the same distribution.
- So Chernoff bounds will hold for the binomial distribution (sum of Bernoulli trials) and for the more general sum of Poisson trials.
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. 
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$\mu = E[X]$$
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$\mu = E[X] = E[\sum_{i=1}^{n} X_i]$$
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$
\mu = E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]
$$
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i$$
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$
\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i
$$

Moment generating functions

For the moment generating functions of each $X_i$:

$$
M_{X_i}(t) = E[e^{tX_i}]
$$
Let \( X_1, \ldots, X_n \) be a sequence of independent Poisson trials with \( P(X_i = 1) = p_i \) and let \( X = \sum_{i=1}^{n} X_i \). Then:

\[
\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i
\]

**Moment generating functions**

For the moment generating functions of each \( X_i \):

\[
M_{X_i}(t) = E[e^{tX_i}] = p_i e^t + (1 - p_i)
\]
Poisson Trials

Expectation

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i$$

Moment generating functions

For the moment generating functions of each $X_i$:

$$M_{X_i}(t) = E[e^{t X_i}] = p_i e^t + (1 - p_i) = 1 + p_i(e^t - 1)$$
Poisson Trials

**Expectation**

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$
\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i
$$

**Moment generating functions**

For the moment generating functions of each $X_i$:

$$
M_{X_i}(t) = E[e^{tX_i}] = p_i e^t + (1 - p_i) = 1 + p_i (e^t - 1) \leq e^{p_i (e^t - 1)}
$$
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i = 1) = p_i$ and let $X = \sum_{i=1}^{n} X_i$. Then:

$$\mu = E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i$$

**Moment generating functions**

For the moment generating functions of each $X_i$:

$$M_{X_i}(t) = E[e^{tX_i}] = p_i e^t + (1 - p_i) = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}$$
Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t - 1)}$$
Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t-1)}$$

But what about $M_X(t)$?
Moment generating functions

Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t - 1)}$$

But what about $M_X(t)$?

Recall that for $X$ and $Y$ independent:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$
Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t-1)}$$

But what about $M_X(t)$?

Recall that for $X$ and $Y$ independent:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Hence:

$$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t)$$
Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t-1)}$$

But what about $M_X(t)$?
Recall that for $X$ and $Y$ independent:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Hence:

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t)$$

$$\leq \prod_{i=1}^n e^{p_i(e^t-1)}$$
Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t - 1)}$$

But what about $M_X(t)$?

Recall that for $X$ and $Y$ independent:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Hence:

$$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$$\leq \prod_{i=1}^{n} e^{p_i(e^t - 1)}$$

$$= e^{\sum_{i=1}^{n} p_i(e^t - 1)}$$
Moment generating functions

Thus, for each $X_i$:

$$M_{X_i}(t) \leq e^{p_i(e^t - 1)}$$

But what about $M_X(t)$?

Recall that for $X$ and $Y$ independent:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

Hence:

$$M_X(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$$\leq \prod_{i=1}^{n} e^{p_i(e^t - 1)}$$

$$= e^{\sum_{i=1}^{n} p_i(e^t - 1)}$$

$$= e^{(e^t - 1)\mu}$$
Example

Using the moment generating function of $X$:
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$. 
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) =$$
Example

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) = \mu e^{(e^t - 1)\mu + t}$$
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) = \mu e^{(e^t - 1)\mu + t}$$

Hence:

$$E[X] =$$
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) = \mu e^{(e^t - 1)\mu + t}$$

Hence:

$$E[X] = M_X^{(1)}(0) \leq$$
Example

Using the moment generating function of $X$:
- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) = \mu e^{(e^t - 1)\mu + t}$$

Hence:

$$E[X] = M_X^{(1)}(0) \leq \mu e^{(e^0 - 1)\mu + 0}$$
Example

Using the moment generating function of $X$:

- Evaluate the expectation of $X$, i.e., $E(X)$.

Answer

For Poisson trials:

$$M_X(t) = e^{(e^t - 1)\mu}$$

The first derivative of the moment generating function of $X$:

$$M_X^{(1)}(t) = \mu e^{(e^t - 1)\mu + t}$$

Hence:

$$E[X] = M_X^{(1)}(0) \leq \mu e^{(e^0 - 1)\mu + 0} = \mu = \sum_{i=1}^{n} p_i$$
Chernoff Bounds

Deriving Chernoff Bounds
We obtain the Chernoff bound for a random variable $X$ by applying Markov's inequality to $e^{tx}$ for some well-chosen value $t$. 
We obtain the Chernoff bound for a random variable $X$ by applying Markov's inequality to $e^{tx}$ for some well-chosen value $t$.

For any $t > 0$, we can derive the following inequality:

$$P(X \geq a) = P(e^{tx} \geq e^{ta})$$
Chernoff Bounds

Deriving Chernoff Bounds

We obtain the Chernoff bound for a random variable $X$ by applying Markov’s inequality to $e^{tx}$ for some well-chosen value $t$.

For any $t > 0$, we can derive the following inequality:

$$P(X \geq a) = P(e^{tx} \geq e^{ta}) \leq \frac{E[e^{tx}]}{e^{ta}}$$
Deriving Chernoff Bounds

We obtain the Chernoff bound for a random variable $X$ by applying Markov's inequality to $e^{tx}$ for some well-chosen value $t$.

For any $t > 0$, we can derive the following inequality:

$$P(X \geq a) = P(e^{tx} \geq e^{ta}) \leq \frac{E[e^{tx}]}{e^{ta}}$$

That is,

$$P(X \geq a) \leq \min_{t>0} \frac{E[e^{tx}]}{e^{ta}}$$
We obtain the Chernoff bound for a random variable $X$ by applying Markov's inequality to $e^{tx}$ for some well-chosen value $t$.

For any $t > 0$, we can derive the following inequality:

$$P(X \geq a) = P(e^{tx} \geq e^{ta}) \leq \frac{E[e^{tx}]}{e^{ta}}$$

That is,

$$P(X \geq a) \leq \min_{t>0} \frac{E[e^{tx}]}{e^{ta}}$$

Similarly, for any $t < 0$:

$$P(X \leq a) \leq \min_{t<0} \frac{E[e^{tx}]}{e^{ta}}$$
Chernoff Bounds

Deriving Chernoff Bounds

We obtain the Chernoff bound for a random variable $X$ by applying Markov's inequality to $e^{tx}$ for some well-chosen value $t$.

For any $t > 0$, we can derive the following inequality:

$$P(X \geq a) = P(e^{tx} \geq e^{ta}) \leq \frac{E[e^{tx}]}{e^{ta}}$$

That is,

$$P(X \geq a) \leq \min_{t>0} \frac{E[e^{tx}]}{e^{ta}}$$

Similarly, for any $t < 0$:

$$P(X \leq a) \leq \min_{t<0} \frac{E[e^{tx}]}{e^{ta}}$$
Chernoff Bounds

Note

- Bounds for specific distributions are obtained by choosing appropriate values for $t$. 
Chernoff Bounds

Note

- Bounds for specific distributions are obtained by choosing appropriate values for $t$.
- Often we choose a value for $t$ that gives convenient bounds (and not the minimum).
Bounds for specific distributions are obtained by choosing appropriate values for $t$. Often we choose a value for $t$ that gives convenient bounds (and not the minimum). Bounds derived this way are (collectively) referred to as *Chernoff bounds*. 

Note
Theorem - Chernoff Bounds

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. 
Theorem - Chernoff Bounds

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then:

for any $\delta > 0$,

$$P(X \geq (1 + \delta)\mu) < \left(\frac{e^{\delta}}{(1 + \delta)(1+\delta)}\right)^{\mu}$$
Theorem - Chernoff Bounds

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then:

- for any $\delta > 0$,
  \[
  P(X \geq (1 + \delta)\mu) < \left( \frac{e^{\delta}}{(1 + \delta)(1+\delta)} \right)^{\mu}
  \]

- for any $0 < \delta \leq 1$,
  \[
  P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\mu \delta^2}{3}}
  \]
Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then:

- for any $\delta > 0$,
  \[
P(X \geq (1 + \delta)\mu) < \left( \frac{e^\delta}{(1 + \delta)(1+\delta)} \right)^\mu
  \]

- for any $0 < \delta \leq 1$,
  \[
P(X \geq (1 + \delta)\mu) \leq e^{-\mu \delta^2/3}
  \]

- for $R \geq 6\mu$,
  \[
P(X \geq R) \leq 2^{-R}
  \]
Theorem - Chernoff Bounds

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then:

- for any $\delta > 0$, 
  $$P(X \geq (1 + \delta)\mu) < \left( \frac{e^{\delta}}{(1 + \delta)^{1+\delta}} \right)^\mu$$

- for any $0 < \delta \leq 1$, 
  $$P(X \geq (1 + \delta)\mu) \leq e^{-\mu \delta^2 / 3}$$

- for $R \geq 6\mu$, 
  $$P(X \geq R) \leq 2^{-R}$$
Chernoff Bounds

Note

- The first bound is the strongest.
Chernoff Bounds

Note

- The first bound is the strongest.
- We derive the other two from the first one.
Chernoff Bounds

Note

- The first bound is the strongest.
- We derive the other two from the first one.
- The other two are easier to compute in many situations.
Chernoff Bounds

Proof.
Proof.

First inequality - we want to show that for any $\delta > 0$:

$$P(X \geq (1 + \delta)\mu) < \left( \frac{e^{\delta}}{(1 + \delta)(1 + \delta)} \right)^\mu$$
Proof.

First inequality - we want to show that for any $\delta > 0$:

$$P(X \geq (1 + \delta)\mu) < \left(\frac{e^{\delta}}{(1 + \delta)^{1+\delta}}\right)^{\mu}$$

For any $t > 0$, by Markov's inequality:

$$P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1+\delta)\mu})$$
Proof.

First inequality - we want to show that for any $\delta > 0$:

$$\Pr(X \geq (1 + \delta)\mu) < \left(\frac{e^\delta}{(1 + \delta)(1+\delta)}\right)^\mu$$

For any $t > 0$, by Markov's inequality:

$$\Pr(X \geq (1 + \delta)\mu) \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}$$

$$= \Pr(e^{tX} \geq e^{t(1+\delta)\mu})$$
Proof.

First inequality - we want to show that for any $\delta > 0$:

$$P(X \geq (1 + \delta)\mu) < \left( \frac{e^{\delta}}{(1 + \delta)^{(1+\delta)}} \right)^\mu$$

For any $t > 0$, by Markov’s inequality:

$$P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1+\delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}$$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$$
Chernoff Bounds

Proof.

First inequality - we want to show that for any $\delta > 0$:

$$P(X \geq (1 + \delta)\mu) < \left( \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^\mu$$

For any $t > 0$, by Markov's inequality:

$$P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1+\delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}$$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$$

For any $\delta > 0$, we can set $t =$
Proof.

First inequality - we want to show that for any \( \delta > 0 \):

\[
P(X \geq (1 + \delta)\mu) < \left( \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^\mu
\]

For any \( t > 0 \), by Markov's inequality:

\[
P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1+\delta)\mu})
\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}
\leq \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}}
\]

For any \( \delta > 0 \), we can set \( t = \ln(1 + \delta) \). Then:
Proof.

First inequality - we want to show that for any $\delta > 0$:

$$P(X \geq (1 + \delta)\mu) < \left( \frac{e^\delta}{(1 + \delta)(1 + \delta)} \right)^\mu$$

For any $t > 0$, by Markov’s inequality:

$$P(X \geq (1 + \delta)\mu) = P(e^{tX} \geq e^{t(1 + \delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1 + \delta)\mu}}$$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1 + \delta)\mu}}$$

For any $\delta > 0$, we can set $t = \ln(1 + \delta)$. Then:

$$P(X \geq (1 + \delta)\mu) < \left( \frac{e^\delta}{(1 + \delta)(1 + \delta)} \right)^\mu$$
Proof.

Second inequality: We want to show that for any $0 < \delta \leq 1$,

$$P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\mu \delta^2}{3}}$$
Proof.

Second inequality: We want to show that for any $0 < \delta \leq 1$,

$$P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\mu \delta^2}{3}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu \leq e^{-\frac{\mu \delta^2}{3}}$$
Proof.

Second inequality: We want to show that for any $0 < \delta \leq 1$,

$$P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\mu\delta^2}{3}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}}\right)^\mu \leq e^{-\frac{\mu\delta^2}{3}}$$

We take the logarithm of both sides:

$$\ln\left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}}\right)^\mu \leq \ln\left(e^{-\frac{\mu\delta^2}{3}}\right)$$
Proof.

Second inequality: We want to show that for any $0 < \delta \leq 1$,

$$P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\mu \delta^2}{3}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^{\delta}}{(1 + \delta)(1+\delta)}\right)^\mu \leq e^{-\frac{\mu \delta^2}{3}}$$

We take the logarithm of both sides:

$$\ln\left(\frac{e^{\delta}}{(1 + \delta)(1+\delta)}\right)^\mu \leq \ln\left(e^{-\frac{\mu \delta^2}{3}}\right)$$

which leads to the following condition (for the second inequality to hold):

$$f(\delta) = \delta - (1 + \delta)\ln(1 + \delta) + \frac{\delta^2}{3} \leq 0$$
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. 
Proof.
We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

$$f'(\delta) = -\ln(1 + \delta) + \frac{2}{3} \delta$$
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3} \delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3} \delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3}\delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
- $f''(\delta) > 0$ for $\delta > \frac{1}{2}$
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3} \delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
- $f''(\delta) > 0$ for $\delta > \frac{1}{2}$

That is, in the interval $[0, 1]$, $f'(\delta)$ first decreases and then increases. But:
Proof.

We want to show that for \(0 < \delta \leq 1\), \(f(\delta) \leq 0\).
First, we compute the derivatives of \(f(\delta)\):

- \(f'(\delta) = -\ln(1 + \delta) + \frac{2}{3}\delta\)
- \(f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}\)

Note that:
- \(f''(\delta) < 0\) for \(0 \leq \delta < \frac{1}{2}\)
- \(f''(\delta) > 0\) for \(\delta > \frac{1}{2}\)

That is, in the interval \([0, 1]\), \(f'(\delta)\) first decreases and then increases. But:
- \(f'(0) = 0\)
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$.

First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3}\delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
- $f''(\delta) > 0$ for $\delta > \frac{1}{2}$

That is, in the interval $[0, 1]$, $f'(\delta)$ first decreases and then increases. But:

- $f'(0) = 0$
- $f'(1) < 0$ (since $\ln 2 > \frac{2}{3}$)
Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3}\delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
- $f''(\delta) > 0$ for $\delta > \frac{1}{2}$

That is, in the interval $[0, 1]$, $f'(\delta)$ first decreases and then increases. But:

- $f'(0) = 0$
- $f'(1) < 0$ (since $\ln 2 > \frac{2}{3}$)

Hence, $f'(\delta) \leq 0$ for $\delta \in [0, 1]$. 
Chernoff Bounds

Proof.

We want to show that for $0 < \delta \leq 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = -\ln(1 + \delta) + \frac{2}{3} \delta$
- $f''(\delta) = -\frac{1}{1+\delta} + \frac{2}{3}$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < \frac{1}{2}$
- $f''(\delta) > 0$ for $\delta > \frac{1}{2}$

That is, in the interval $[0, 1]$, $f'(\delta)$ first decreases and then increases. But:

- $f'(0) = 0$
- $f'(1) < 0$ (since $\ln 2 > \frac{2}{3}$)

Hence, $f'(\delta) \leq 0$ for $\delta \in [0, 1]$.
Since $f(0) = 0$, $f(\delta) \leq 0$ for $0 < \delta \leq 1$. 

□
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$. 
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.
Then, for $R \geq 6\mu$, we have $\delta \geq 5$. 
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.

Then, for $R \geq 6\mu$, we have $\delta \geq 5$.

Using the first inequality, we have:

$$P(X \geq R) = P(X \geq (1 + \delta)\mu) \leq \left( \frac{e^\delta}{(1 + \delta)(1+\delta)} \right)^\mu$$
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.

Then, for $R \geq 6\mu$, we have $\delta \geq 5$.

Using the first inequality, we have:

$$P(X \geq R) = P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1 + \delta)(1 + \delta)}\right)^\mu$$

$$\leq \left(\frac{e}{1 + \delta}\right)^{(1+\delta)\mu}$$
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.
Then, for $R \geq 6\mu$, we have $\delta \geq 5$.

Using the first inequality, we have:

$$P(X \geq R) = P(X \geq (1 + \delta)\mu) \leq \left( \frac{e^\delta}{(1 + \delta)(1+\delta)} \right)^\mu$$

$$\leq \left( \frac{e}{1 + \delta} \right)^{(1+\delta)\mu}$$

$$= \left( \frac{e}{1 + \delta} \right)^R$$
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.

Then, for $R \geq 6\mu$, we have $\delta \geq 5$.

Using the first inequality, we have:

$$P(X \geq R) = P(X \geq (1 + \delta)\mu) \leq \left( \frac{e^{\delta}}{(1 + \delta)^{(1+\delta)}} \right)^{\mu}$$

$$\leq \left( \frac{e}{1 + \delta} \right)^{(1+\delta)\mu}$$

$$= \left( \frac{e}{1 + \delta} \right)^{R}$$

$$\leq \left( \frac{e}{6} \right)^{R}$$
Proof.

Third inequality - we want to show that for $R \geq 6\mu$:

$$P(X \geq R) \leq 2^{-R}$$

We set $R = (1 + \delta)\mu$.
Then, for $R \geq 6\mu$, we have $\delta \geq 5$.

Using the first inequality, we have:

$$P(X \geq R) = P(X \geq (1 + \delta)\mu) \leq \left( \frac{e^\delta}{(1 + \delta)(1 + \delta)} \right)^\mu$$

$$\leq \left( \frac{e}{1 + \delta} \right)^{(1 + \delta)\mu}$$

$$= \left( \frac{e}{1 + \delta} \right)^R$$

$$\leq \left( \frac{e}{6} \right)^R$$

$$\leq 2^{-R}$$

Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$: 


Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^n X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)(1-\delta)}\right)^\mu$$
Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

1. $P(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1 - \delta)(1 - \delta)} \right)^\mu$
2. $P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$
Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

- $P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)(1-\delta)}\right)^{\mu}$
- $P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$

Note

- Again, the first bound is stronger.
**Theorem - Chernoff Bounds - Deviation below the mean**

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

- $P(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1-\delta)(1-\delta)} \right)^\mu$

- $P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu \delta^2}{2}}$

**Note**

- Again, the first bound is stronger.
- The second is derived from the first.
Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

1. $P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)(1-\delta)}\right)^{\mu}$

2. $P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$

Note

- Again, the first bound is stronger.
- The second is derived from the first.
- The second is generally easier to use and sufficient.
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}}\right)^\mu$$
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1 - \delta)}\right)^\mu$$

Again, we have to properly choose $t$. 

Chernoff Bounds
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1 - \delta)}\right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov’s inequality:

$$P(X \leq (1 - \delta)\mu)$$
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1-\delta)}\right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov’s inequality:

$$P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1-\delta)\mu})$$
Proof.

First inequality - we want to show that for \(0 < \delta < 1\):

\[
P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1-\delta)}\right)^\mu
\]

Again, we have to properly choose \(t\). For any \(t < 0\), by Markov’s inequality:

\[
P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1-\delta)\mu}) \\ \leq \frac{E[e^{tX}]}{e^{t(1-\delta)\mu}}
\]
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1 - \delta)(1 - \delta)} \right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov’s inequality:

$$P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1 - \delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1 - \delta)\mu}}$$

$$\leq \frac{e^{t(\epsilon - 1)\mu}}{e^{t(1 - \delta)\mu}}$$
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}}\right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov’s inequality:

$$P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1-\delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1-\delta)\mu}}$$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu}}$$

For $0 < \delta < 1$, we can set $t =$
Chernoff Bounds

Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1 - \delta)(1-\delta)} \right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov's inequality:

$$P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1-\delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1-\delta)\mu}}$$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu}}$$

For $0 < \delta < 1$, we can set $t = \ln(1 - \delta)$. Then:
Proof.

First inequality - we want to show that for $0 < \delta < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1 - \delta)}\right)^\mu$$

Again, we have to properly choose $t$. For any $t < 0$, by Markov's inequality:

$$P(X \leq (1 - \delta)\mu) = P(e^{tX} \geq e^{t(1 - \delta)\mu}) \leq \frac{E[e^{tX}]}{e^{t(1 - \delta)\mu}} \leq \frac{e^{(e^t - 1)\mu}}{e^{t(1 - \delta)\mu}}$$

For $0 < \delta < 1$, we can set $t = \ln(1 - \delta)$. Then:

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)(1 - \delta)}\right)^\mu$$
Proof.

Second inequality: We want to show that for any $0 < \delta < 1$,

$$P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$$
Proof.

Second inequality: We want to show that for any $0 < \delta < 1$,

$$P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^{-\delta}}{(1 - \delta)(1-\delta)}\right)^\mu \leq e^{-\frac{\mu\delta^2}{2}}$$
Chernoff Bounds

Proof.

Second inequality: We want to show that for any $0 < \delta < 1$,

$$P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}}\right)^\mu \leq e^{-\frac{\mu\delta^2}{2}} \quad \text{or} \quad \frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \leq e^{-\frac{\delta^2}{2}}$$

We take the logarithm of both sides:

$$\ln \frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \leq \ln\left(e^{-\frac{\delta^2}{2}}\right)$$
Proof.

Second inequality: We want to show that for any $0 < \delta < 1$,

$$P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$$

Hence, with respect to the first inequality, we want to show:

$$\left(\frac{e^{-\delta}}{(1 - \delta)(1-\delta)}\right)^\mu \leq e^{-\frac{\mu\delta^2}{2}} \text{ or } \frac{e^{-\delta}}{(1 - \delta)(1-\delta)} \leq e^{-\frac{\delta^2}{2}}$$

We take the logarithm of both sides:

$$\ln \frac{e^{-\delta}}{(1 - \delta)(1-\delta)} \leq \ln \left( e^{-\frac{\delta^2}{2}} \right)$$

which leads to the following condition (for the second inequality to hold):

$$f(\delta) = -\delta - (1 - \delta)\ln(1 - \delta) + \frac{\delta^2}{2} \leq 0$$
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. 
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:
- $f''(\delta) < 0$ for $0 \leq \delta < 1$
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < 1$

That is, in the interval $[0, 1]$, $f'(\delta)$ decreases. But:
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < 1$

That is, in the interval $[0, 1]$, $f'(\delta)$ decreases. But:

- $f'(0) = 0$
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < 1$

That is, in the interval $[0,1]$, $f'(\delta)$ decreases. But:

- $f'(0) = 0$

Hence, $f'(\delta) \leq 0$ for $\delta \in [0,1)$.
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$. First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:
- $f''(\delta) < 0$ for $0 \leq \delta < 1$

That is, in the interval $[0, 1]$, $f'(\delta)$ decreases. But:
- $f'(0) = 0$

Hence, $f'(\delta) \leq 0$ for $\delta \in [0, 1)$. So, $f(\delta)$ is non-increasing for $\delta \in [0, 1)$. 
Proof.

Again, we want to show that for $0 < \delta < 1$, $f(\delta) \leq 0$.
First, we compute the derivatives of $f(\delta)$:

- $f'(\delta) = \ln(1 - \delta) + \delta$
- $f''(\delta) = -\frac{1}{1-\delta} + 1$

Note that:

- $f''(\delta) < 0$ for $0 \leq \delta < 1$

That is, in the interval $[0, 1]$, $f'(\delta)$ decreases. But:

- $f'(0) = 0$

Hence, $f'(\delta) \leq 0$ for $\delta \in [0, 1)$.
So, $f(\delta)$ is non-increasing for $\delta \in [0, 1)$.
Since $f(0) = 0$, $f(\delta) \leq 0$ for $0 < \delta < 1$. 

\[\square\]
Corollary

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$: 
Corollary

Let $X_1, \ldots, X_n$ be a sequence of independent Poisson trials with $P(X_i) = p_i$, $X = \sum_{i=1}^{n} X_i$, and $\mu = E[X]$. Then, for $0 < \delta < 1$:

$$P(|X - \mu| \geq \delta \mu) \leq 2e^{-\frac{\mu \delta^2}{3}}$$