Chernoff Bounds (Fundamentals)

Pavlos Eirinakis[1]

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

14 February, 2012

Outline

1. Tail bounds

Outline

Outline

Outline

Tail bounds

Tail bounds

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.

Tail bounds

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.

Tail bounds

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.

Tail bounds

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 $\frac{3}{4} \cdot n$?

Tail bounds

Markov's inequality

Tail bounds

Markov's inequality

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

Tail bounds

Markov's inequality

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

$$P(X \geq c) \leq \frac{E(X)}{c}$$

Tail bounds

Markov's inequality

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

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

Tail bounds

Markov's inequality

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

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

Chebyshev's inequality

Tail bounds

Markov's inequality

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

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

Chebyshev's inequality

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

Tail bounds

Markov's inequality

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

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

Chebyshev's inequality

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

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

Tail bounds

Markov's inequality

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

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

Chebyshev's inequality

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

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

Tail bounds

Markov's inequality

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

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

Chebyshev's inequality

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

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

Example

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

Moment Generating Functions

The moment-generating function of a random variable X

Moment Generating Functions

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.

Moment Generating Functions

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:

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

Moment Generating Functions

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 $M_X(0) = 1$.

Moment Generating Functions

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 $M_X(0) = 1$.

Moment Generating Functions

Theorem

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

Moment Generating Functions

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,

Moment Generating Functions

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 $t = 0$.

Moment Generating Functions

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. □

Moment Generating Functions

Bernoulli random variables

Consider a Bernoulli random variable X.

Moment Generating Functions

Bernoulli random variables

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

Moment Generating Functions

Bernoulli random variables

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

$$M_X(t) \quad =$$

Moment Generating Functions

Bernoulli random variables

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

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

Moment Generating Functions

Bernoulli random variables

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

$$
\begin{aligned}
M_X(t) &= E[e^{tX}] \\
&= \sum_X P(X)e^{tX}
\end{aligned}
$$

Moment Generating Functions

Bernoulli random variables

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

$$\begin{aligned} M_X(t) &= E[e^{tX}] \\ &= \sum_X P(X)e^{tX} \\ &= P(X=1)e^t + P(X=0)e^0 \end{aligned}$$

Moment Generating Functions

Bernoulli random variables

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

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

Moment Generating Functions

Bernoulli random variables

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

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

Example

Using the moment generating function of X:

Moment Generating Functions

Bernoulli random variables

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

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

Example

Using the moment generating function of X:

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

Moment Generating Functions

Bernoulli random variables

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

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

Example

Using the moment generating function of X:

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

Moment Generating Functions

Bernoulli random variables

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

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

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

Moment Generating Functions

Geometric random variables

Consider a geometric random variable X with parameter p.

Moment Generating Functions

Geometric random variables

Consider a geometric random variable X with parameter p.
What is a geometric random variable?

Moment Generating Functions

Geometric random variables

Consider a geometric random variable X with parameter p.
What is a geometric random variable?
What is its probability distribution?

Moment Generating Functions

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

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

Moment Generating Functions

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:

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

Moment Generating Functions

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:

Moment Generating Functions

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:

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

Moment Generating Functions

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:

$$\begin{aligned} M_X(t) &= E[e^{tX}] \\ &= \frac{p}{1-p}((1-(1-p)e^t)^{-1} - 1). \end{aligned}$$

Moment Generating Functions

Geometric random variables

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

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

Moment Generating Functions

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:

Moment Generating Functions

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., $E(X)$.

Moment Generating Functions

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., $Var(X)$.

Moment Generating Functions

Geometric random variables

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

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

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:

Moment Generating Functions

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 $t < -\ln(1-p)$:

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

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:

$$\begin{aligned} M_X^{(1)}(t) &= p(1-(1-p)e^t)^{-2}e^t \\ M_X^{(2)}(t) &= 2p(1-p)(1-(1-p)e^t)^{-3}e^{2t} + p(1-(1-p)e^t)^{-2}e^t \end{aligned}$$

Moment Generating Functions

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:

$$\begin{aligned} M_X^{(1)}(t) &= p(1-(1-p)e^t)^{-2}e^t \\ M_X^{(2)}(t) &= 2p(1-p)(1-(1-p)e^t)^{-3}e^{2t} + p(1-(1-p)e^t)^{-2}e^t \end{aligned}$$

Moment Generating Functions

Theorem

Moment Generating Functions

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.

Moment Generating Functions

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

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

Moment Generating Functions

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.

Moment Generating Functions

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.

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

Moment Generating Functions

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.

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

Moment Generating Functions

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.

$$\begin{aligned} M_{X+Y}(t) &= E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] \\ &= E[e^{tX}]E[e^{tY}] \end{aligned}$$

Moment Generating Functions

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.

$$\begin{aligned} M_{X+Y}(t) &= E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] \\ &= E[e^{tX}]E[e^{tY}] \\ &= M_X(t)M_Y(t) \end{aligned}$$

□

Poisson Trials

Poisson Trials

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

Poisson Trials

Poisson Trials

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

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.

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.
- The distribution of the random variables in Poisson trials are not necessarily identical.

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.
- 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.

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.
- 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.

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 $X = \sum_{i=1}^{n} X_i$.

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:

$$\mu = E[X]$$

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:

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

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:

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

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:

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

Poisson Trials

Expectation

Let $X_1, \dots, X_n$ 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:

$$M_{X_i}(t) = E[e^{tX_i}]$$

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:

$$\begin{aligned} M_{X_i}(t) &= E[e^{tX_i}] \\ &= p_i e^t + (1 - p_i) \end{aligned}$$

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:

$$\begin{aligned} M_{X_i}(t) &= E[e^{tX_i}] \\ &= p_i e^t + (1 - p_i) \\ &= 1 + p_i(e^t - 1) \end{aligned}$$

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:

$$\begin{aligned} M_{X_i}(t) &= E[e^{tX_i}] \\ &= p_i e^t + (1 - p_i) \\ &= 1 + p_i(e^t - 1) \\ &\leq e^{p_i(e^t - 1)} \end{aligned}$$

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:

$$\begin{aligned} M_{X_i}(t) &= E[e^{tX_i}] \\ &= p_i e^t + (1 - p_i) \\ &= 1 + p_i(e^t - 1) \\ &\leq e^{p_i(e^t - 1)} \end{aligned}$$

Poisson Trials

Moment generating functions

Thus, for each X_i:

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

Poisson Trials

Moment generating functions

Thus, for each X_i:

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

But what about $M_X(t)$?

Poisson Trials

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:

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

Poisson Trials

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:

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

Poisson Trials

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:

$$\begin{aligned} M_X(t) &= \Pi_{i=1}^{n} M_{X_i}(t) \\ &\leq \Pi_{i=1}^{n} e^{p_i(e^t-1)} \end{aligned}$$

Poisson Trials

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:

$$\begin{aligned} M_X(t) &= \Pi_{i=1}^{n} M_{X_i}(t) \\ &\leq \Pi_{i=1}^{n} e^{p_i(e^t-1)} \\ &= e^{\sum_{i=1}^{n} p_i(e^t-1)} \end{aligned}$$

Poisson Trials

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:

$$\begin{aligned} M_X(t) &= \Pi_{i=1}^n M_{X_i}(t) \\ &\leq \Pi_{i=1}^n e^{p_i(e^t-1)} \\ &= e^{\sum_{i=1}^n p_i(e^t-1)} \\ &= e^{(e^t-1)\mu} \end{aligned}$$

Poisson Trials

Example

Using the moment generating function of X:

Poisson Trials

Example

Using the moment generating function of X:

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

Poisson Trials

Example

Using the moment generating function of X:

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

Answer

For Poisson trials:

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

Poisson Trials

Example

Using the moment generating function of X:

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

Answer

For Poisson trials:

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

The first derivative of the moment generating function of X:

$$M_X^{(1)}(t) \quad =$$

Poisson Trials

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:

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

Poisson Trials

Example

Using the moment generating function of X:

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

Answer

For Poisson trials:

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

The first derivative of the moment generating function of X:

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

Hence:

$$E[X] =$$

Poisson Trials

Example

Using the moment generating function of X:

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

Answer

For Poisson trials:

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

The first derivative of the moment generating function of X:

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

Hence:

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

Poisson Trials

Example

Using the moment generating function of X:

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

Answer

For Poisson trials:

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

The first derivative of the moment generating function of X:

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

Hence:

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

Poisson Trials

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

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.

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:

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

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:

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

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,

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

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

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).

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 derived this way are (collectively) referred to as *Chernoff bounds*.

Chernoff Bounds

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 $\mu = E[X]$.

Chernoff Bounds

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:

1. for any $\delta > 0$,

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

Chernoff Bounds

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:

1. for any $\delta > 0$,

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

2. for any $0 < \delta \leq 1$,

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

Chernoff Bounds

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:

1. for any $\delta > 0$,
$$P(X \geq (1+\delta)\mu) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$
2. for any $0 < \delta \leq 1$,
$$P(X \geq (1+\delta)\mu) \leq e^{-\frac{\mu\delta^2}{3}}$$
3. for $R \geq 6\mu$,
$$P(X \geq R) \leq 2^{-R}$$

Chernoff Bounds

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:

1. for any $\delta > 0$,

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

2. for any $0 < \delta \leq 1$,

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

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.

Chernoff Bounds

Proof.

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

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

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:

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

Chernoff Bounds

Proof.

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

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

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

$$\begin{aligned} P(X \geq (1+\delta)\mu) &= P(e^{tX} \geq e^{t(1+\delta)\mu}) \\ &\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}} \end{aligned}$$

Chernoff Bounds

Proof.

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

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

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

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

Chernoff Bounds

Proof.

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

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

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

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

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

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:

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

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

Chernoff Bounds

Proof.

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

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

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

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

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

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

□

Chernoff Bounds

Proof.

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

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

Chernoff Bounds

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:

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

Chernoff Bounds

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:

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

Chernoff Bounds

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:

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

We take the logarithm of both sides:

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

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

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

Chernoff Bounds

Proof.

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

Chernoff Bounds

Proof.

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

Chernoff Bounds

Proof.

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

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

Chernoff Bounds

Proof.

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

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

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 $0 \leq \delta < \frac{1}{2}$

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 $\delta > \frac{1}{2}$

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:

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:

- $f'(0) = 0$

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 $\ln 2 > \frac{2}{3}$)

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 $\delta \in [0,1]$.

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 $0 < \delta \leq 1$.

□

Chernoff Bounds

Proof.

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

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

Chernoff Bounds

Proof.

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

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

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

Chernoff Bounds

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 $\delta \geq 5$.

Chernoff Bounds

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:

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

Chernoff Bounds

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:

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

Chernoff Bounds

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:

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

Chernoff Bounds

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:

$$\begin{aligned} P(X \geq R) = P(X \geq (1+\delta)\mu) &\leq (\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}})^{\mu} \\ &\leq (\frac{e}{1+\delta})^{(1+\delta)\mu} \\ &= (\frac{e}{1+\delta})^{R} \\ &\leq (\frac{e}{6})^{R} \end{aligned}$$

Chernoff Bounds

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:

$$\begin{aligned} P(X \geq R) = P(X \geq (1+\delta)\mu) &\leq (\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}})^{\mu} \\ &\leq (\frac{e}{1+\delta})^{(1+\delta)\mu} \\ &= (\frac{e}{1+\delta})^{R} \\ &\leq (\frac{e}{6})^{R} \\ &\leq 2^{-R} \end{aligned}$$

□

Chernoff Bounds

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 $0 < \delta < 1$:

Chernoff Bounds

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

Chernoff Bounds

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

Chernoff Bounds

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

Note

- Again, the first bound is stronger.

Chernoff Bounds

Theorem - Chernoff Bounds - Deviation below the mean

Let $X_1, \dots, X_n$ 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.

Chernoff Bounds

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

Chernoff Bounds

Proof.

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

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

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.

Chernoff Bounds

Proof.

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

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

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

$$P(X \leq (1-\delta)\mu) \quad =$$

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:

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

Chernoff Bounds

Proof.

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

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

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

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

Chernoff Bounds

Proof.

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

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

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

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

Chernoff Bounds

Proof.

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

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

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

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

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:

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

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

Chernoff Bounds

Proof.

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

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

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

$$\begin{aligned} P(X \le (1-\delta)\mu) &= P(e^{tX} \ge e^{t(1-\delta)\mu)}) \\ &\le \frac{E[e^{tX}]}{e^{t(1-\delta)\mu)}} \\ &\le \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu)}} \end{aligned}$$

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

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

□

Chernoff Bounds

Proof.

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

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

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:

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

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:

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

We take the logarithm of both sides:

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

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:

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

We take the logarithm of both sides:

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

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

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

Chernoff Bounds

Proof.

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

Chernoff Bounds

Proof.

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

Chernoff Bounds

Proof.

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

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

Chernoff Bounds

Proof.

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

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

Chernoff Bounds

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 $0 \leq \delta < 1$

Chernoff Bounds

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:

Chernoff Bounds

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:

- $f'(0) = 0$

Chernoff Bounds

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 $\delta \in [0,1)$.

Chernoff Bounds

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 $\delta \in [0,1)$.

Chernoff Bounds

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 $0 < \delta < 1$. □

Chernoff Bounds

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 $0 < \delta < 1$:

Chernoff Bounds

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 $0 < \delta < 1$:

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