RAD (4) Tail inequalities, chernoff, set balancing

Fremlæggelse

Poisson trials & Poisson Binomial Distribution

Let $0 \leq p_1, \ldots, p_n \leq 1$, let $X_1, \ldots, X_n$ be independent indicator variables with $\Pr[X_i = 1] = p_i$, and let $X = \sum_{i=1}^{n} X_i$. We call $X_1, \ldots X_n$ Poisson Trials, and say that X has the Poisson Binomial Distribution.

Bernoulli trials

Let $0 \leq p \leq 1$, let $X_1, \ldots, X_n$ be independent indicator variables with $\Pr[X_i = 1] = p$, and let $X = \sum_{i=1}^{n} X_i$. We call $X_1, \ldots X_n$ Bernoulli Trials, and say that X has the Binomial Distribution.

First Chernoff Bound

Given a random variable X with the Poisson Binomial Distribution:

- For $\delta > 0$, find small $\epsilon > 0$ so that

  $$\Pr[X > (1+\delta)\mu] < \epsilon$$

  Let $X_1, \ldots, X_n$ be independent Poisson trials such that for $1 \leq i \leq n, \Pr[X_i = 1] = p_i$, where $0 < p_i < 1$. Let $X = \sum_{i=1}^{n} X_i$ and $\mu \geq \mathbb{E}[X] = \sum_{i=1}^{n} p_i$. For any $\delta > 0$,

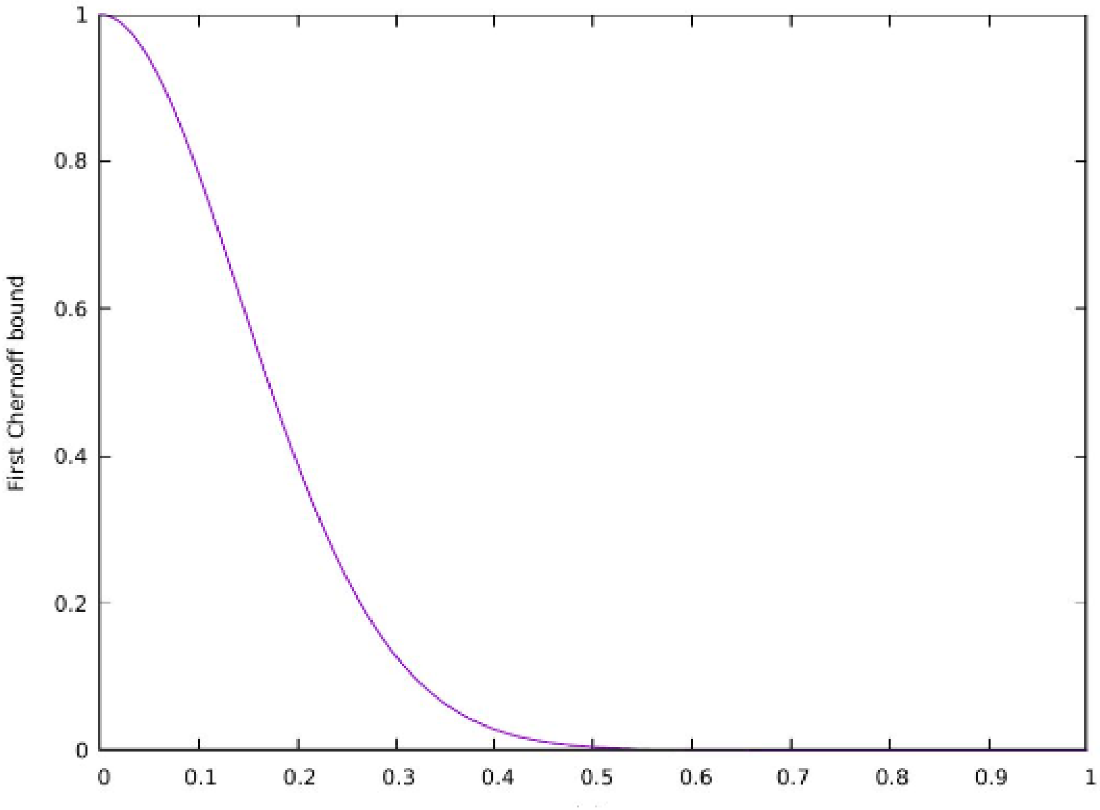
  $$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} < e^{-\frac{\delta^2}{3}\mu}$$

  $e^{-\frac{\delta^2}{3}\mu}$ er ikke en del af pensum at vise.

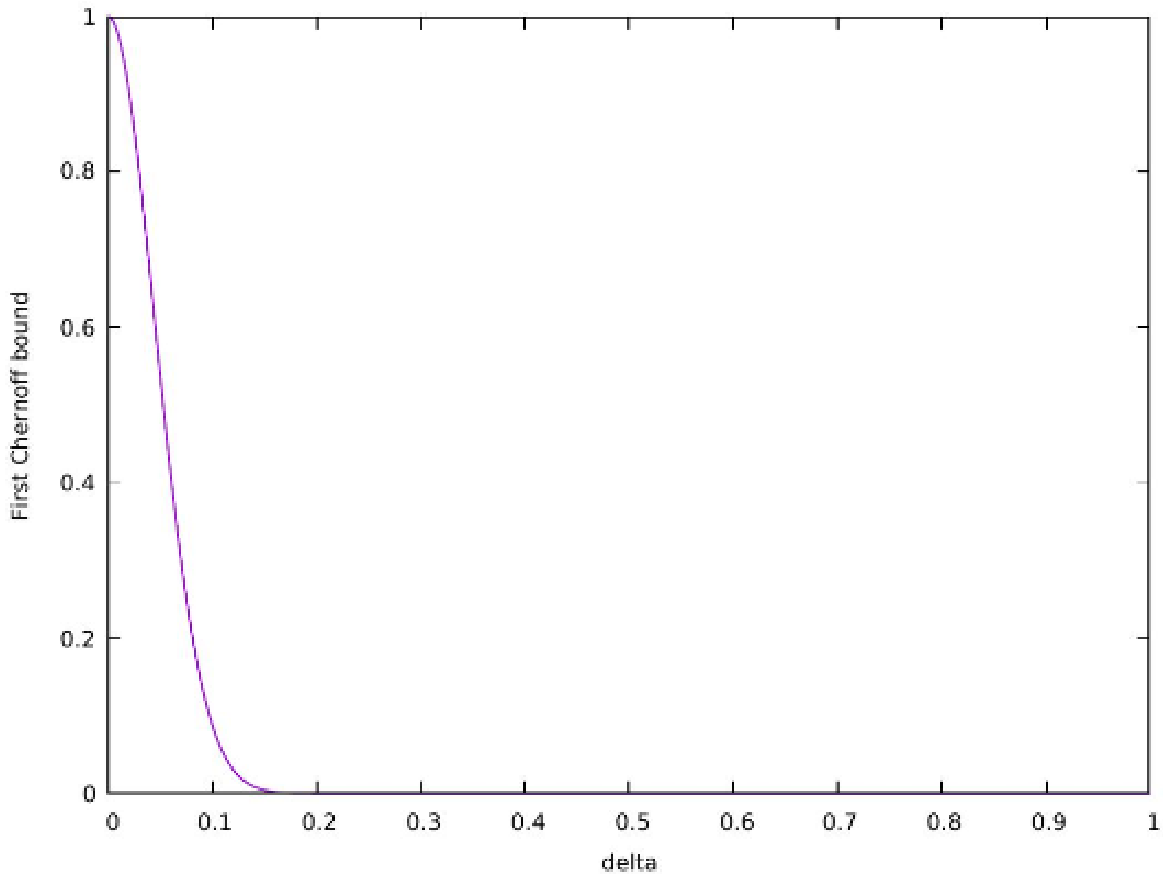
Consider n independent tosses of a fair coin and let X denote the number of heads. For $\frac{1}{2} < q \leq 1$, which δ should we choose to upper bound $\Pr[X > qn]$?

We have $\mu = n/2$ so $(1+\delta)n/2 = qn \Leftrightarrow \delta = 2q - 1$.

Example with $n = 100$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$:



Example with $n = 1000$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$:



A slightly weaker bound is

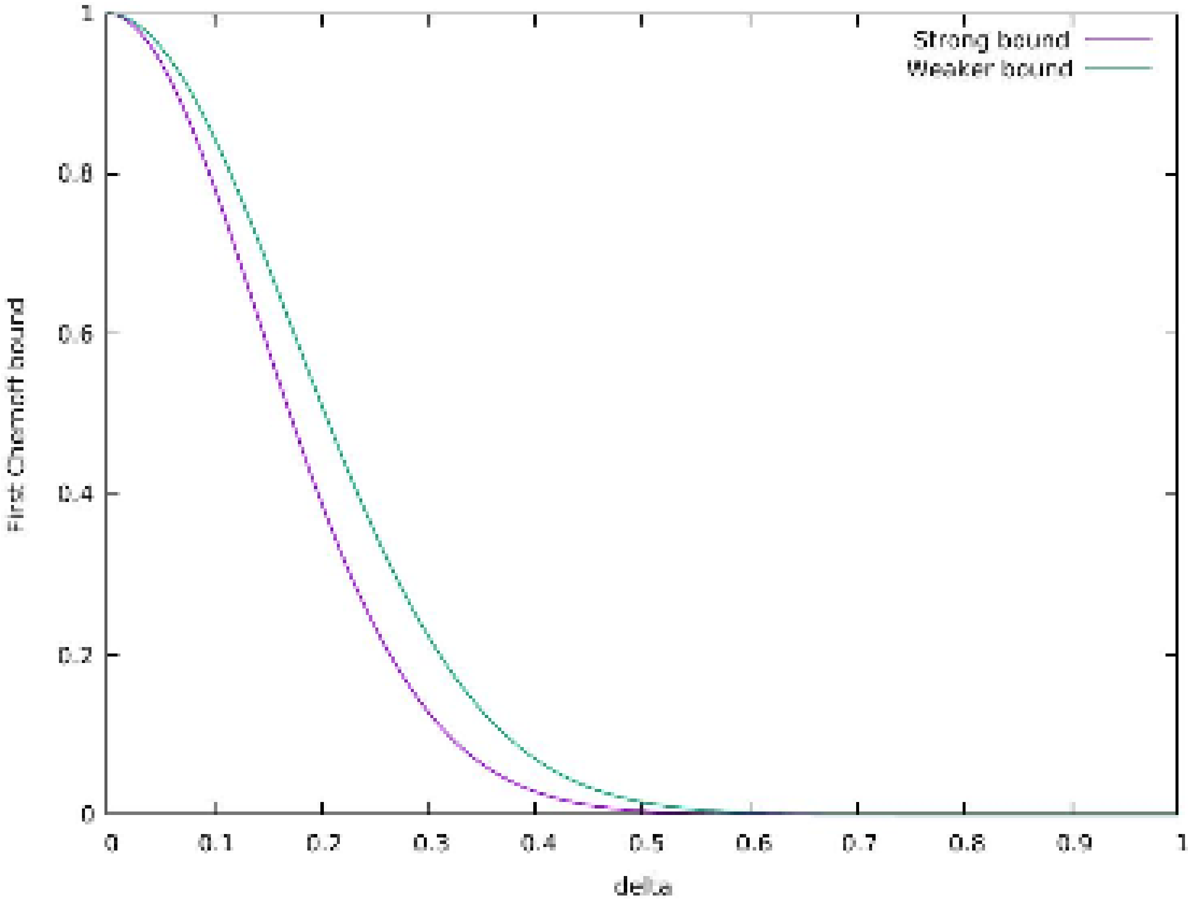
$$\Pr[X > (1+\delta)\mu] < e^{-\frac{\delta^2}{2+\delta}\mu}$$

When $0 < \delta \leq 1$, an even weaker bound is:

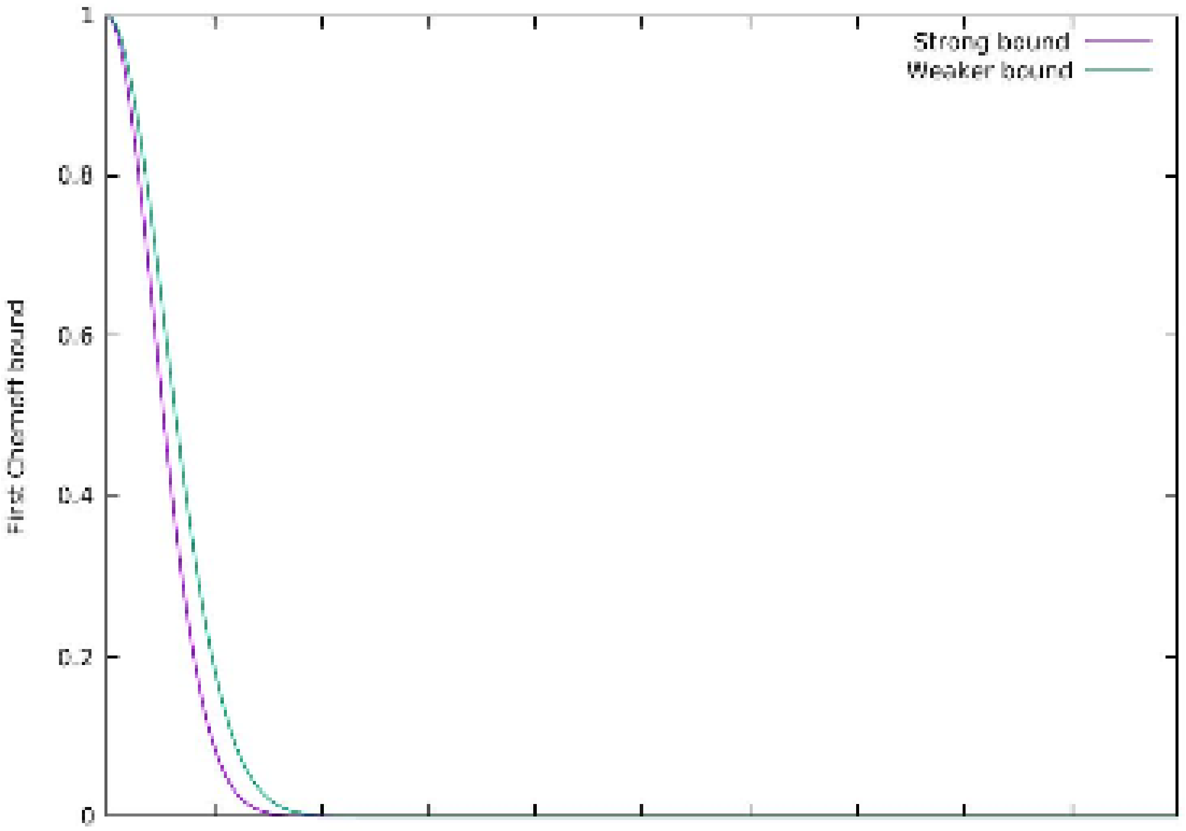
$$\Pr[X > (1+\delta)\mu] < e^{-\frac{\delta^2}{3}\mu}$$

These bounds are useful since they are often easier to work with in proofs. We will not prove these weaker bounds.

Example with $n = 100$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$ and weaker bound $e^{-\frac{\delta^2}{3}\mu}$:



Example with $n = 1000$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$ and weaker bound $e^{-\frac{\delta^2}{3}\mu}$:



Proof

We will use the following lemma:
Let $Y_1, \ldots, Y_k$ be independent variables. Then

$$\mathbb{E}\left[\prod_{i=1}^{k} Y_i\right] = \prod_{i=1}^{k} \mathbb{E}\left[Y_i\right]$$

Main ideas:
Given a random variable X with the Poisson Binomial Distribution:

- For $\delta > 0$, find small $\epsilon > 0$ so that

  $$\Pr[X > (1+\delta)\mu] < \epsilon$$

  Let $X_1, \ldots, X_n$ be independent Poisson trials such that for $1 \leq i \leq n, \Pr\left[X_i = 1\right] = p_i$, where $0 < p_i < 1$. Let $X = \sum_{i=1}^{n} X_i$ and $\mu \geq \mathbb{E}[X] = \sum_{i=1}^{n} p_i$. For any $\delta > 0$,
- Analyze $(1+\delta)^X$ rather than X.
- Apply Markov's inequality to $(1+\delta)^X$.
- Use independence to turn expectation of a product into a product of expectations.

vi gør basen til en potens og vælger $1+\delta$ til at være en ny base.
vi kan f.eks. omskrive $5 > 2$ til $3^5 > 3^2$ og så holder det stadig. det er det samme vi gør.

$$\Pr[X > (1+\delta)\mu] = \Pr\left[(1+\delta)^x > (1+\delta)^{(1+\delta)\mu}\right]$$

Markov's første ulighed fordi x > 0. Her er $x = (1+\delta)^x$,
$t = (1+\delta)^{(1+\delta)\mu}$.

$$< \frac{\mathbb{E}\left[(1+\delta)^x\right]}{(1+\delta)^{(1+\delta)\mu}}$$

Vi ganger $1+\delta$ med sig selv x gange, derfor kan vi bruge product af expectation når de er uafhængige.

$\mathbb{E}\left[(1+\delta)^x\right] = \mathbb{E}\left[(1+\delta)^{\sum_{i=1}^n x_i}\right] = \mathbb{E}\left[\prod_{i=1}^n (1+\delta)^{x_i}\right] = \prod_{i=1}^n \mathbb{E}\left[(1+\delta)^{x_i}\right]$

$$= \frac{\prod_{i=1}^n \mathbb{E}\left[(1+\delta)^{x_i}\right]}{(1+\delta)^{(1+\delta)\mu}}$$

Det er en indikator variabel så den kan kun være 0 eller 1.

$\mathbb{E}\left[(1+\delta)^{x_i}\right] = (1-p_i)(1+\delta)^0 + p_i(1+\delta)^1 = 1 + p_i\delta$

der gælder at forventningen af en indikatorvariabel er sansyndligheden for at den forekommer.

$$= \frac{\prod_{i=1}^n (1+p_i\delta)}{(1+\delta)^{(1+\delta)\mu}}$$

bruger at $1+x \leq e^x$.

$$\leq \frac{\prod_{i=1}^n e^{p_i\delta}}{(1+\delta)^{(1+\delta)\mu}}$$

potens regnereglen $x^a \cdot x^b = x^{a+b}$

$$= \frac{e^{\left(\sum_{i=1}^n p_i\delta\right)}}{(1+\delta)^{(1+\delta)\mu}}$$

$e^{\left(\sum_{i=1}^n p_i\delta\right)} \leq e^{\delta\mu}$ fordi $e^\delta > 1$ and $\mu \geq \sum_{i=1}^n p_i$

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

trækker μ ud.

Derved har vi $\Pr[X > (1+\delta)\mu] < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$

Second Chernoff Bound

- For $0 < \delta < 1$, find small $\epsilon > 0$ so that

$$\Pr[X < (1-\delta)\mu] < \epsilon$$

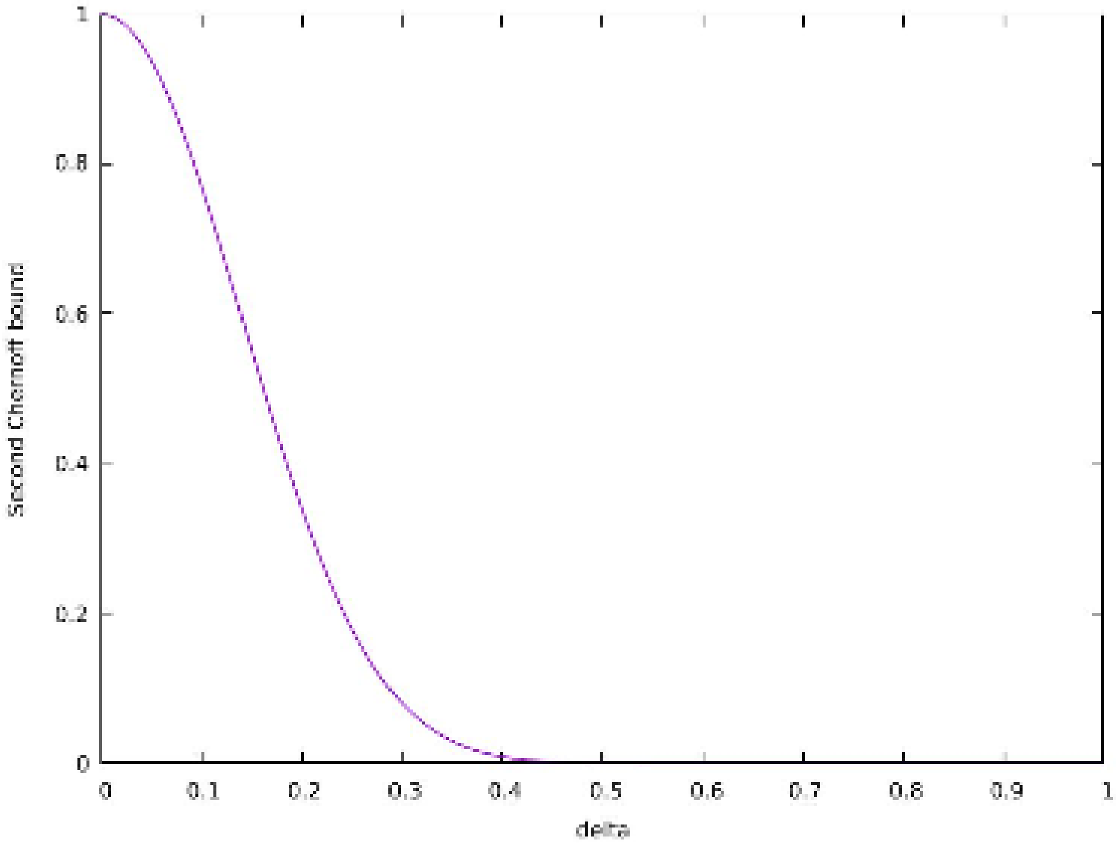
Let $X_1, \ldots, X_n$ be independent Poisson trials such that, for $1 \leq i \leq n, \Pr\left[X_i = 1\right] = p_i$, where $0 < p_i < 1$. Let $X = \sum_{i=1}^n X_i$ and $\mu \leq \mathbb{E}[X] = \sum_{i=1}^n p_i$. For any $0 < \delta < 1$,

$$\begin{aligned}\Pr[X < (1-\delta)\mu] &< \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \\ &< e^{-\frac{\delta^2}{2}\mu} (\text{ Theorem 4.2})\end{aligned}$$

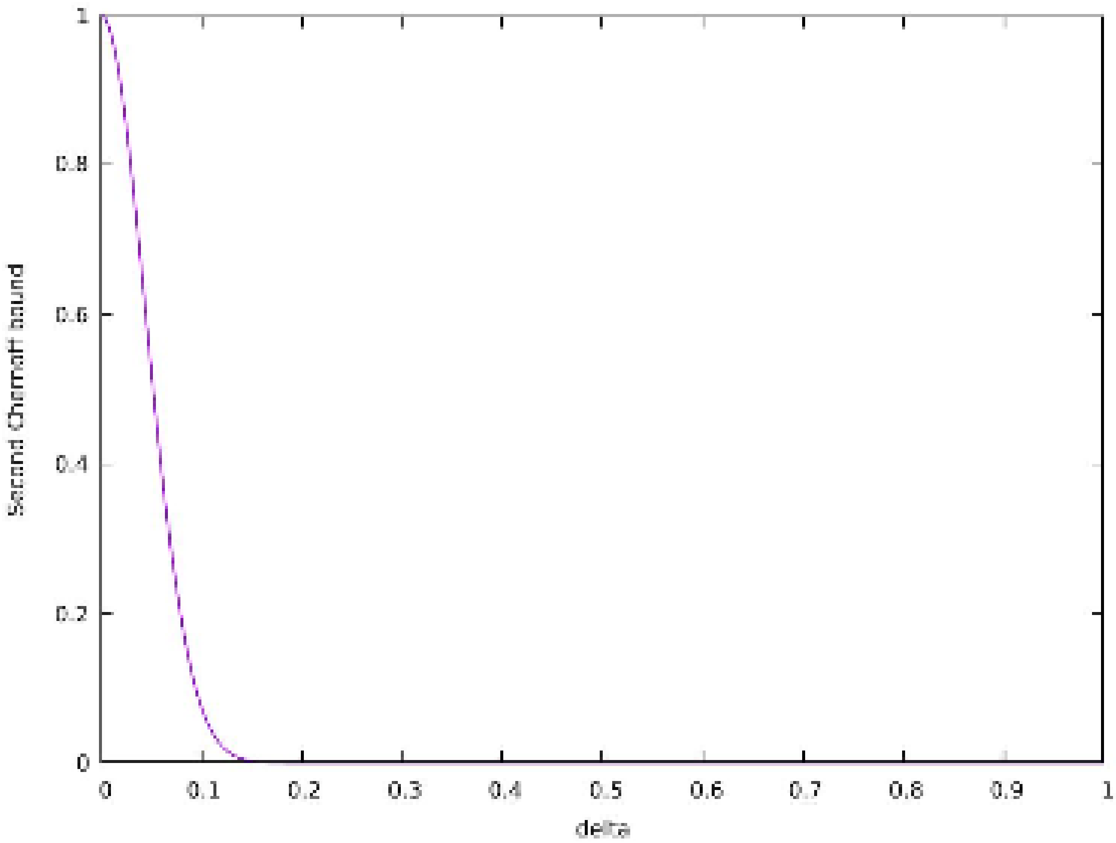
This is exactly the same setup as for the first Chernoff bound, except that we require $\mu \leq \mathbb{E}[X]$ instead of $\mu \geq \mathbb{E}[X]$. Compare with the previous bound: for any $\delta > 0$,

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

Example with $n = 100$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$:



Example with $n = 1000$ and $p_i = \frac{1}{2}$ for $1 \leq i \leq n$:



Proof

beviset er identisk men det er $1-\delta$ istedet og uligheden er vendt i $\Pr[X<(1-\delta)\mu]$ ellers er alle skridt ens.

$$\begin{aligned}
\Pr[X<(1-\delta)\mu] &= \Pr\left[(1-\delta)^x > (1-\delta)^{(1-\delta)\mu}\right] \\
&< \frac{\mathbb{E}\left[(1-\delta)^x\right]}{(1-\delta)^{(1-\delta)\mu}} \\
&= \frac{\prod_{i=1}^n \mathbb{E}\left[(1-\delta)^{x_i}\right]}{(1-\delta)^{(1-\delta)\mu}} \\
&= \frac{\prod_{i=1}^n (1-p_i\delta)}{(1-\delta)^{(1-\delta)\mu}} \\
&\le \frac{\prod_{i=1}^n e^{-p_i\delta}}{(1-\delta)^{(1-\delta)\mu}} \\
&= \frac{e^{-\delta\sum_{i=1}^n p_i}}{(1-\delta)^{(1-\delta)\mu}} \\
&\le \frac{e^{-\delta\mu}}{(1-\delta)^{(1-\delta)\mu}} = \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}
\end{aligned}$$

Set balancing

en algoritme der finder en B vektor der kan ganges på A for at minimere max normen. Det kan blandt andet bruges til udvælgelse af forsøgs personer i kliniske forsøg.

$$\underbrace{\begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}\pm 1\\ \pm 1\\ \pm 1\end{bmatrix}}_{B} = AB$$

$$|AB|_\infty$$

Ved at bruge Chernoff kan vi opnå en øvre grænse for fejl sandsynligheden. Det betyder altså at vi ikke behøver at sandsynligheden for algoritmen fejler er meget lav. Vi behøver i fleste tilfælde ikke at køre

$$< 2\sqrt{2n\ln(n)}$$

Chernoff in set balancing