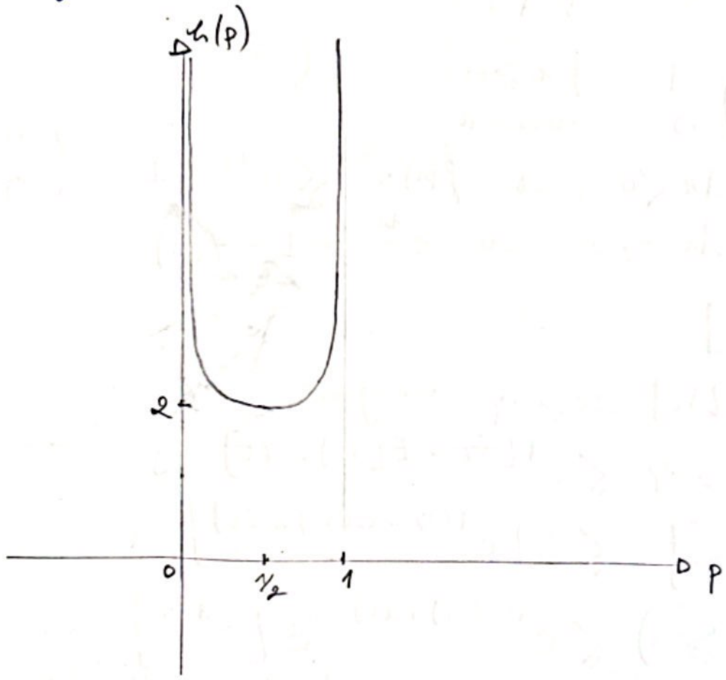
Hoeffding's inequality for a sum of Random variables:

SARA EL KHBIR

- Let $X_1, \ldots, X_\ell$ be independant RV such as $0 \leq X_i \leq 1$.
- Let $S_\ell = \sum_{i=1}^{\ell} X_i$ and $p = E\left[\frac{S_\ell}{\ell}\right]$

1) we have $h(p) = \begin{cases} \frac{1}{1-2p} \log\left(\frac{1-p}{p}\right) & \text{if } 0 < p < \frac{1}{2} \\ \frac{1}{2p(1-p)} & \text{if } \frac{1}{2} \leq p < 1 \end{cases}$

the graph of this function is:



→ For $0 < p < \frac{1}{2}$, $h(p)$ increases as p moves towards 0. this means that the probability of the event S_ℓ exceeding $(p+t)\ell$ is more sensitive to changes in p when p is small (towards 1/2)

And for $\frac{1}{2} \leq p < 1$, it also means that the event above is more sensitive to changes towards (1/2).

→ the inequality quantifies how rare it is for S_ℓ to significantly exceed or fall below p by a certain threshold t as ℓ become large.

2) by convexity of the $e^{\lambda x}$ function ($\lambda \geq 0$)

we have $\forall\, a \leq x \leq b$:

$$e^{\lambda x} \leq \frac{x-a}{b-a} e^{\lambda b} + \frac{b-x}{b-a} e^{\lambda a}$$

thus by monotonicity of the expectation, we have for X a RV such as $a \leq X \leq b$

$$E[e^{\lambda X}] \leq E\left[\frac{X-a}{b-a}\right] e^{\lambda b} + E\left[\frac{b-X}{b-a}\right] e^{\lambda a}$$

$$\Rightarrow \quad E[e^{\lambda X}] \leq \frac{E(X)-a}{b-a} e^{\lambda b} + \frac{b-E(X)}{b-a} e^{\lambda a}$$

3). Proof that $\mathbb{1}_{x\geq 0} \leq e^{\lambda x} \quad \forall \lambda > 0$

$\rightarrow$ let $f(x) = \mathbb{1}_{x\geq 0} = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$

let $x \in \mathbb{R}$: • if $x < 0 \Rightarrow \lambda x < 0 \Rightarrow f(x) = 0 \leq e^{\lambda x} \leq 1$ (positivity of the exponential function

• if $x \geq 0 \Rightarrow \lambda x \geq 0 \Rightarrow e^{\lambda x} \geq 1 = f(x)$

so $\boxed{\mathbb{1}_{x\geq 0} \leq e^{\lambda x}}$

Applying this for $\mathbb{1}\{S_\ell - E[S_\ell] - \ell t \geq 0\}$ we get

$$\mathbb{1}\{S_\ell - E[S_\ell] - \ell t \geq 0\} \leq e^{\lambda[S_\ell - E[S_\ell] - \ell t]}$$

$$\Rightarrow \quad E\left[\mathbb{1}\{S_\ell - E(S_\ell) - \ell t \geq 0\}\right] \leq E\left[e^{\lambda(S_\ell - E(S_\ell) - \ell t)}\right]$$

$$\Rightarrow \quad P\left(S_\ell - E(S_\ell) - \ell t \geq 0\right) \leq e^{-\lambda(E(S_\ell) + \ell t)} E\left[e^{\lambda S_\ell}\right]$$

$$\Rightarrow \quad P\left(S_\ell \geq \ell(p+t)\right) \leq e^{-\lambda(p+t)\ell} E\left[e^{\lambda \sum_{i=1}^{\ell} x_i}\right] = e^{-\lambda(p+t)\ell} E\left[\sum_{i=1}^{\ell} \lambda x_i\right]$$

$$\Rightarrow \quad P\left(S_\ell \geq \ell(p+t)\right) \leq e^{-\lambda(p+t)\ell} \prod_{i=1}^{\ell} E(e^{\lambda x_i})$$ (by independence of the RV x_i)

4) let $p_i = E[x_i]$,

applying question 2 with $a=0$ and $b=1$, we have $\forall i \in [\![1, \ell]\!]$:

$$E[e^{\lambda x_i}] \leq 1 - E(x_i) + e^{\lambda} E(x_i)$$

$$\Rightarrow \quad E(e^{\lambda x_i}) \leq 1 - p_i + p_i e^{\lambda}.$$

By non negativity of the expectation we have :

$$\prod_{i=1}^{\ell} E(e^{\lambda x_i}) \leq \prod_{i=1}^{\ell} (1 - p_i + e^{\lambda} p_i)$$

5) Proof of the geometric arithmetic mean inequality:

→ let $a_1 — a_\ell$ positive real numbers:

- we have $\forall x \in \mathbb{R}$: $\ln(x) \leq x - 1$.

(this inequality follows from the function $f(x) = \ln x - x$ having a maximum at 1 and being concave. ($f'(x) = \frac{1}{x} - 1$ and $f''(x) = -\frac{1}{x^2} \leq 0$))

Now let $A = \frac{\sum_{i=1}^{\ell} a_i}{\ell}$, using the inequality, we get $\forall i \in [\![1..\ell]\!]$:

$$\ln\left(\frac{a_i}{A}\right) \leq \frac{a_i}{A} - 1$$

and summing these inequalities $\forall i$, we get: $\sum_{i=1}^{\ell} \ln\left(\frac{a_i}{A}\right) \leq \sum_{i=1}^{\ell} \frac{a_i}{A} - \ell$

$$\Rightarrow \ln\left(\frac{\prod_{i=1}^{\ell} a_i}{A^\ell}\right) \leq \underbrace{\frac{\sum a_i}{A}}_{=\ell} - \ell$$

$$\Rightarrow \ln\left(\frac{\prod_{i=1}^{\ell} a_i}{A^\ell}\right) \leq 0. \Rightarrow \frac{\prod_{i=1}^{\ell} a_i}{A^\ell} \leq 1$$

(by setting exponential to both sides)

$$\Rightarrow \left(\prod_{i=1}^{\ell} a_i\right)^{1/\ell} \leq A \Rightarrow \boxed{\left(\prod_{i=1}^{\ell} a_i\right)^{1/\ell} \leq \frac{1}{\ell} \sum_{i=1}^{\ell} a_i}$$

6) we have from question 4:

$$\prod_{i=1}^{\ell} E(e^{\lambda X_i}) \leq \prod_{i=1}^{\ell} (1 - p_i + p_i e^\lambda)$$

as $1 - p_i + p_i e^\lambda \geq 0$, we have $\prod_{i=1}^{\ell} (1 - p_i + p_i e^\lambda) \leq \left(\frac{1}{\ell} \sum_{i=1}^{\ell} 1 - p_i + p_i e^\lambda\right)^\ell$

and $\left(\frac{1}{\ell} \sum_{i=1}^{\ell} 1 - p_i + p_i e^\lambda\right)^\ell = \left(\frac{1}{\ell}\left(\ell - \sum_{i=1}^{\ell} E(X_i) + e^\lambda \sum_{i=1}^{\ell} E(X_i)\right)\right)^\ell = \left(1 - p + e^\lambda p\right)^\ell$

so $\boxed{\prod_{i=1}^{\ell} E(e^{\lambda X_i}) \leq (1 - p + e^\lambda p)^\ell}$

7) From question 3, we have $P(S_\ell \geq (p+t)\ell) \leq e^{-\lambda(p+t)\ell} \prod_{i=1}^{\ell} E(e^{\lambda X_i})$

and from question 6, we have:

$$P(S_\ell \geq (p+t)\ell) \leq e^{-\lambda(p+t)\ell} (1 - p + e^\lambda p)^\ell$$

→ let $f(\lambda) = e^{-\lambda(p+t)\ell} (1 - p + e^\lambda p)^\ell$

we have $f'(\lambda) = -(p+t)\ell\, e^{-\lambda(p+t)\ell} (1 - p + e^\lambda p)^\ell + e^{-\lambda(p+t)\ell} \times \ell \times p (1 - p + p e^\lambda)^{\ell-1}$

so $f'(\lambda) = 0 \iff -\ell(p+t)(1 - p + e^\lambda p) + \ell p = 0$

$\iff e^\lambda = \frac{1}{p}\left(\frac{\ell p}{\ell(p+t)} - 1 + p\right) \iff \lambda = \log \frac{(1-p)(p+t)}{(1-p-t)p}$.

In order to have: $\lambda = \log\left(\frac{(1-p)(p+t)}{(1-p-t)p}\right)$ positive, we should have $\frac{(1-p)(p+t)}{(1-p-t)p} \geq 1$

which is $\Leftrightarrow (1-p)(p+t) \geq (1-p-t)p \Leftrightarrow p-p^2+t-pt \geq p-p^2-pt$

$\Leftrightarrow t \geq 0$ which is true.

Now taking this value for λ in the inequality, we have:

$$P(S_\ell \geq (p+t)\ell) \leqslant e^{-(p+t)\ell \log\left(\frac{(1-p)(p+t)}{(1-p-t)p}\right)} \left[1-p+p\left(\frac{(1-p)(p+t)}{p(1-p-t)}\right)\right]^{\ell}$$

$$= \left[\frac{(1-p)(p+t)}{(1-p-t)p}\right]^{-\ell(p+t)} \left[(1-p)\left(\frac{1}{1-p-t}\right)\right]^{\ell}$$

$$= \left(\frac{p}{p+t}\right)^{\ell(p+t)} \left(\frac{1-p-t}{1-p}\right)^{\ell(p+t)} \left(\frac{1-p}{1-p-t}\right)^{-\ell}$$

so we conclude:

$$\boxed{P(S_\ell \geq (p+t)\ell) \leqslant \left(\frac{p}{p+t}\right)^{\ell(p+t)} \left(\frac{1-p}{1-p-t}\right)^{\ell(1-p-t)}}$$

8) We have: $G(t,p) = \frac{p+t}{t^2}\log\left(\frac{p+t}{p}\right) + \frac{1-p-t}{t^2}\log\left(\frac{1-p-t}{1-p}\right)$

$$\Rightarrow \frac{\partial G(t,p)}{\partial t} = \frac{t^2-2t(p+t)}{t^4}\log\left(\frac{p+t}{p}\right) + \frac{p+t}{t^2}\,\frac{1/p}{1+t/p} + \frac{-t^2-2t(1-p-t)}{t^4}\log\left(\frac{1-p-t}{1-p}\right)$$

$$+ \frac{1-p-t}{t^2}\,\frac{-1/1-p}{1-t/1-p}.$$

$$\Rightarrow t^2\frac{\partial G(t,p)}{\partial t} = \frac{-(t+2p)}{t}\log\left(\frac{p+t}{p}\right) + \cancel{1} + \frac{t-2+2p}{t}\log\left(\frac{1-p-t}{1-p}\right) - \cancel{1}$$

$$= \left(1-2\frac{1-p}{t}\right)\log\left(1-\frac{t}{1-p}\right) - \left(1-2\frac{p+t}{t}\right)\log\left(1-\frac{t}{t+p}\right)$$

setting $H(x) = \left(1-\frac{2}{x}\right)\log(1-x)$, we get:

$$\boxed{t^2\frac{\partial}{\partial t}G(t,p) = H\left(\frac{t}{1-p}\right) - H\left(\frac{t}{t+p}\right)}$$

→ we have $H(x) = \left(1-\frac{2}{x}\right)\log(1-x)$

By the formula of Taylor Young expansion of usual functions

$$\log(1-x) \underset{x\to 0}{=} -x - \frac{x^2}{2} - \cdots - \frac{x^n}{n} + o(x^n)$$

So $H(x) = \left(1-\frac{2}{x}\right)\left(-x-\frac{x^2}{2} - \cdots - \frac{x^n}{n} + o(x^n)\right)$

$$= -\cancel{x} - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} \cdots + 2 + \cancel{x} + \frac{2}{3}x^2 + \frac{2}{4}x^3 + \frac{2}{5}x^4 + \cdots$$

$$= 2 + \left(\frac{2}{3}-\frac{1}{2}\right)x^2 + \left(\frac{2}{4}-\frac{1}{3}\right)x^3 + \left(\frac{2}{5}-\frac{1}{4}\right)x^4 + \cdots$$

we have $H'(x) = 2\left(\frac{2}{3}-\frac{1}{2}\right)x + 3\left(\frac{2}{4}-\frac{1}{3}\right)x^2 + 4\left(\frac{2}{5}-\frac{1}{4}\right)x^3 + \cdots$

which is positive $\forall x$ such as $0 < x < 1$ so $H(x)$ is increasing in $]0,1[$

$\Rightarrow$ As $t^2 \frac{\partial}{\partial t} G(t,p) = H\left(\frac{t}{1-p}\right) - H\left(\frac{t}{t+p}\right)$

and $H(x)$ is increasing in $]0,1[$. $\left(\begin{array}{l}\text{we have } p>0 \Rightarrow 0<\frac{t}{t+p}<1 \\ \text{and } t<1-p \Rightarrow 0<\frac{t}{1-p}<1\end{array}\right)$

so $\frac{\partial G}{\partial t}(t,p) > 0 \iff H\left(\frac{t}{1-p}\right) - H\left(\frac{t}{t+p}\right) > 0$

$$\iff \frac{t}{1-p} - \frac{t}{t+p} > 0 \iff \frac{1}{1-p} > \frac{1}{t+p} \iff \boxed{t > 1-2p}$$

$\rightarrow$ When $1-2p>0$:

$G(t,p)$ attain its minimum $\iff \frac{\partial}{\partial t} G(t,p) = 0$

$$\iff H\left(\frac{t}{1-p}\right) = H\left(\frac{t}{t+p}\right) \underset{\substack{H \text{ strictly} \\ \text{increasing in }]0,1[}}{\iff} \frac{t}{1-p} = \frac{t}{t+p} \iff \boxed{t = 1-2p}$$

with the corresponding minimal value being:

$$G(1-2p,p) = \frac{p+(1-2p)}{(1-2p)^2} \log\left(\frac{p+(1-2p)}{p}\right) + \frac{1-p-(1-2p)}{(1-2p)^2} \log\left(1 - \frac{1-2p}{1-2p+p}\right)$$

$$= \frac{1-p}{(1-2p)^2} \log\left(\frac{-p+1}{p}\right) + \frac{p}{(1-2p)^2} \log\left(\frac{p}{1-p}\right)$$

$$= \left(\frac{1}{1-2p}\right) \log\left(\frac{1-p}{p}\right) = h(p).$$

g) if $1-2p \leq 0$ then $G(t,p)$ attain its minimum when $t \to 0$.

In such case:

$\rightarrow$ we have $G(t,p) = \frac{p+t}{t^2} \log\left(1+\frac{t}{p}\right) + \frac{1-p-t}{t^2} \log\left(1 - \frac{t}{1-p}\right)$

using the taylor expansion of $\log(1+t)$ around $t=0$: we get:

$$G(t,p) = \frac{p+t}{t^2}\left(\frac{t}{p} - \frac{t^2}{2p^2} + \frac{t^3}{3p^3} + \cdots\right) + \frac{1-p-t}{t^2}\left(\frac{-t}{1-p} - \frac{t^2}{2(1-p)^2} - \frac{t^3}{3(1-p)^3} + \cdots\right)$$

expanding up to $o(t^3)$, we get:

$$G(t,p) = \frac{p+t}{pt} - \frac{p+t}{2p^2} + \frac{(p+t)t}{3p^3} - \frac{(1-p-t)}{(1-p)t} - \frac{(1-p-t)}{2(1-p)^2} - \frac{t(1-p-t)}{3(1-p)^3}$$

$$= \left[\frac{p+t}{pt} - \frac{(1-p-t)}{(1-p)t}\right] - \left[\frac{p+t}{2p^2} - \frac{(1-p-t)}{2(1-p)^2}\right] + t\left[\frac{p+t}{3p^3} - \frac{(1-p-t)}{3(1-p)^3}\right]$$

$$= \frac{1}{p(1-p)} - \left(\frac{p+t}{2p^2} - \frac{1-p-t}{2(1-p)^2}\right) + t\left(\frac{p+t}{3p^3} - \frac{(1-p-t)}{3(1-p)^3}\right)$$

$$\xrightarrow[t\to 0]{} \frac{1}{p(1-p)} - \left(\frac{1}{2p} - \frac{1}{2(1-p)}\right) = \frac{1}{p(1-p)} - \frac{1}{2}\left(\frac{1}{p(1-p)}\right) = \frac{1}{2(p(1-p))}$$

So $\boxed{\lim_{t \to 0} G(t,p) = \dfrac{1}{2p(1-p)} = h(p)}$

$\Rightarrow$ we deduce that $h(p) \leq \inf\limits_{0<t\leq 1-p} G(t,p)$

So we get the second part of the inequality:

$$\boxed{P(S_\ell \geq (p+t)\ell) \leq \left(\frac{p}{p+t}\right)^{\ell(p+t)} \left(\frac{1-p}{1-p-t}\right)^{\ell(1-p-t)} \leq e^{-\ell t^2 h(p)}}$$

$\Rightarrow$ Finally, as $h(p) = \begin{cases} \dfrac{1}{1-2p} \log\left(\dfrac{1-p}{p}\right) & \text{if } 0 < p < \frac{1}{2} \\ \dfrac{1}{2p(1-p)} & \text{if } \frac{1}{2} \leq p < 1 \end{cases}$

- for $\frac{1}{2} \leq p \leq 1$, we have $h'(p) = \dfrac{2p-1}{2p^2(1-p)^2} \geq 0 \Rightarrow h(p)$ increases in $[1/2, 1[$

  and $h(p)$ is minimal for $p = \frac{1}{2} \Rightarrow h(p) \geq h(1/2) = 2$

- for $0 < p < \frac{1}{2}$:

$$h'(p) = \frac{1}{(1-2p)^2}\left[2\log\left(\frac{1-p}{p}\right) - \frac{(1-2p)}{p(1-p)}\right] \leq 0$$

$\Rightarrow h(p)$ is decreasing in $]0, \frac{1}{2}[$

and $h(p) \geq \lim\limits_{p \to \frac{1}{2}} h(p) = 2.$

$\Rightarrow \boxed{h(p) \geq h(1/2) = 2}$

So we get the final ~~right~~ right part of the inequality, and:

$$\boxed{P(S_\ell \geq (p+t)\ell) \leq \left(\frac{p}{p+t}\right)^{\ell(p+t)} \left(\frac{1-p}{1-p-t}\right)^{\ell(1-p-t)} \leq e^{-\ell t^2 h(p)} \leq e^{-2\ell t^2}}$$

Scanned with CamScanner