Tail Bounds

For a random variable X, the tails of X are the parts of the f_X that are "far" from the mean.

MARKOV'S INEQUALITY

Thm: Let be X a non negative variable (i.e. $\mathbb{P}(X>0)=1$), $\forall \xi > 0$:

1) $\mathbb{P}(X > \xi) \leq \frac{\mathbb{E}(X)}{\xi}$

2) $\mathbb{P}(X > \xi\, \mathbb{E}X) \leq \frac{1}{\xi}$

3) $\mathbb{P}(X > \xi) \leq \frac{\mathbb{E}(\varphi(X))}{\varphi(\xi)}$ where φ is a nondecreasing non negative function

CHEBYSHEV'S INEQUALITY

Thm: Let be X an integrable r.v. with $\mathrm{Var}X \neq 0$, $\forall \xi > 0$

1) $\mathbb{P}(|X - \mathbb{E}X| \geq \xi \cdot \mathrm{Var}X) \leq \frac{1}{\xi^2}$

2) $\mathbb{P}(|X - \mathbb{E}X| \geq \xi) \leq \frac{\mathrm{Var}\,X}{\xi^2}$

CHERNOFF BOUNDS

Given X, the moment generating function is $M_X(t) = \mathbb{E}(e^{tX})$

Thm: Let X be a random variable, then $\forall \xi \in \mathbb{R}$

if $t>0$ $\quad \mathbb{P}(X \geq \xi) \leq e^{-\xi t} \cdot M_X(t)$

if $t<0$ $\quad \mathbb{P}(X \leq \xi) \leq e^{-\xi t} \cdot M_X(t)$

where $M_X(\cdot)$ is the moment generating function

HOEFFDING INEQUALITY

Lemma : Hoeffding

Let be Y r.v. such that $Y \in [a,b]$ a.s., then

$$\mathbb{E}\left[e^{t(Y-\mathbb{E}Y)}\right] \leq e^{\frac{t^2(b-a)^2}{8}}$$

Thm: Let $X_1, \dots, X_n$ i.r.v. such that $\forall i = 1, \dots, n$, $a_i \leq X_i \leq b_i$ almost surely. Let $S_n = X_1 + \dots + X_n$, then $\forall \xi > 0$

$$\mathbb{P}(S_n - \mathbb{E}(S_n) > \xi) \leq e^{-\frac{2\xi^2}{\sum (b_i - a_i)^2}}$$

Lemma : Popoviciu inequality

Let be Y r.v. such that $Y \in [a,b]$ a.s., then

$$\mathrm{Var}(Y) \leq \frac{(b-a)^2}{4}$$

BENNETT'S INEQUALITY

Thm: Let $X_1, \dots, X_n$ i.r.v. such that $\forall i = 1, \dots, n$,

(a) $\mathrm{Var}\,X_i$ exists finite

(b) $|X_i - \mathbb{E}X_i| \leq M_i$ a.s.

Let $S_n := X_1 + \dots + X_n$, $M = \max\{M_i\}$, then $\forall \xi \geq 0$

$$\mathbb{P}(S_n - \mathbb{E}S_n > \xi) \leq e^{-\frac{\mathrm{Var}\,S_n}{M^2} \cdot L\left(\frac{M\xi}{\mathrm{Var}\,S_n}\right)}$$

where $L(u) := (1+u)\cdot \log(1+u) - u$

BERNSTEIN'S INEQUALITY

Thm: Let $X_1, \dots, X_n$ i.r.v. such that $\forall i = 1, \dots, n$, $a_i \leq X_i \leq b_i$ almost surely. Let $S_n = X_1 + \dots + X_n$, and $M = \max\{b_i\}$

Then $\forall \xi > 0$

$$\mathbb{P}(S_n - \mathbb{E}(S_n) > \xi) \leq e^{-\frac{\xi^2}{2} \cdot \frac{1}{\sum \mathrm{Var}\,X_i + \frac{1}{3}M\cdot\xi}}$$

CHERNOFF BOUND for binomial

Thm: Let X be a random variable, $X \sim Bin(n,p)$, $\forall \xi > 0$

$$\mathbb{P}(X - \mathbb{E}X \geq \xi) \leq e^{-\frac{2\xi^2}{n}}$$

$$\mathbb{P}(X - \mathbb{E}X \leq -\xi) \leq e^{-\frac{2\xi^2}{n}}$$

STRONGER CHERNOFF BOUND for binomial

Lem: Let $X \sim Bin(n,p)$, then:

$$M_X(t) = \left(p \cdot e^t + 1 - p\right)^n \leq e^{\mathbb{E}X \cdot (e^t - 1)}$$

Thm: Let $X \sim Bin(n,p)$, $\forall \xi > 0$

$$\mathbb{P}(X - \mathbb{E}X \geq \xi \cdot \mathbb{E}X) \leq \left(\frac{e^{\xi}}{(1+\xi)^{1+\xi}}\right)^{\mathbb{E}X}$$

Futhermore, when $0 < \xi < 1$

$$\mathbb{P}(X - \mathbb{E}X \leq -\xi \cdot \mathbb{E}X) \leq \left(\frac{e^{-\xi}}{(1-\xi)^{1-\xi}}\right)^{\mathbb{E}X}$$

Tail Bounds

MARKOV'S INEQUALITY

Thm: Let be X a non negative variable (i.e. $\mathbb{P}(X>0)=1$), $\forall \xi > 0$:

1) $\mathbb{P}(X > \xi) \leq \dfrac{\mathbb{E}(X)}{\xi}$

2) $\mathbb{P}(X > \xi\,\mathbb{E}X) \leq \dfrac{1}{\xi}$

3) $\mathbb{P}(X > \xi) \leq \dfrac{\mathbb{E}(\varphi(X))}{\varphi(\xi)}$ where φ is a nondecreasing non negative function

PROOF: Since $\mathbb{P}(X>0)=1$, $X : \mathbb{R} \longrightarrow [0,+\infty)$. If $\xi \in [0,1]$, clear, $\xi > 1$

Let be X a continuous random variable, then

☆ $\mathbb{E}X = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx = \int_{-\infty}^{0} x \cdot f(x)\,dx + \int_{0}^{+\infty} x \cdot f(x)\,dx =$

$= 0 + \int_{0}^{+\infty} x\, f(x)\,dx$

1) $\mathbb{E}X = \int_{0}^{+\infty} x\, f(x)\,dx =$

$= \int_{0}^{\xi} x\, f(x)\,dx + \int_{\xi}^{+\infty} x f(x)\,dx \geq$

$= \int_{\xi}^{+\infty} x\, f(x)\,dx \geq \xi \int_{\xi}^{+\infty} f(x) \cdot dx = \xi\, \mathbb{P}(X \geq \xi)$

2) Let $\xi' := \mathbb{E}X \cdot \xi > 0$, since $\mathbb{E}X \geq 0$, then

$\mathbb{P}(X \geq \xi') = \mathbb{P}(X \geq \xi \cdot \mathbb{E}X) \underset{(1)}{\leq} \dfrac{1}{\xi} = \dfrac{\mathbb{E}X}{\xi'}$

3) $\mathbb{E}(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(x) \cdot f(x)\,dx =$

$= \int_{-\infty}^{\xi} \varphi(x) \cdot f(x)\,dx + \int_{\xi}^{+\infty} \varphi(x) \cdot f(x)\,dx =$

$= \int_{\xi}^{+\infty} \varphi(x) \cdot f(x)\,dx =$

$\underset{\varphi \text{ is not descending}}{=} \varphi(\xi) \cdot \int_{\xi}^{+\infty} f(x)\,dx$

Let be X a discrete random variable, then

1) $\mathbb{E}X = \sum_{x \in \Omega} x \cdot \mathbb{P}(X=x) =$

$= \sum_{\substack{x \in \Omega \\ x \leq \xi}} x \cdot \mathbb{P}(X=x) + \sum_{\substack{x \in \Omega \\ x > \xi}} x \cdot \mathbb{P}(X=x) \geq$

$\geq \xi \sum_{\substack{x \in \Omega \\ x > \xi}} \mathbb{P}(X=x) = \xi\, \mathbb{P}(X > \xi)$

2) Let $\xi' := \mathbb{E}X \cdot \xi > 0$, since $\mathbb{E}X \geq 0$, then

$\mathbb{P}(X > \xi') = \mathbb{P}(X > \xi \cdot \mathbb{E}X) \underset{(1)}{\leq} \dfrac{1}{\xi} = \dfrac{\mathbb{E}X}{\xi'}$

3) $\mathbb{E}(\varphi(X)) = \sum_{x \in \Omega} \varphi(x) \cdot \mathbb{P}(X=x) =$

$= \sum_{\substack{x \in \Omega \\ x \leq \xi}} \varphi(x) \cdot \mathbb{P}(X=x) + \sum_{\substack{x \in \Omega \\ x > \xi}} \varphi(x) \cdot \mathbb{P}(X=x) \geq$

$\geq \varphi(\xi) \cdot \sum_{\substack{x \in \Omega \\ x > \xi}} \mathbb{P}(X=x) = \varphi(\xi) \cdot \mathbb{P}(X > \xi)$ $\blacksquare$

CHEBYSHEV'S INEQUALITY

Thm: Let be X an integrable r.v. with $\mathrm{Var}\,X \neq 0$, $\forall \xi > 0$

1) $\mathbb{P}\left(|X - \mathbb{E}X| \geq \xi\sqrt{\mathrm{Var}X}\right) \leq \frac{1}{\xi^2}$

2) $\mathbb{P}\left(|X - \mathbb{E}X| \geq \xi\right) \leq \frac{\mathrm{Var}\,X}{\xi^2}$

PROOF: Knowing that $\mathrm{Var}\,X = \mathbb{E}\left[(X - \mathbb{E}X)^2\right]$, we can apply Markov inequality to $(X-\mathbb{E}X)^2$ that is a positive r.v.
Hence we get:

$$\mathbb{P}\left[(X-\mathbb{E}X)^2 \geq \xi^2\right] \leq \frac{\mathbb{E}[(X-\mathbb{E}X)^2]}{\xi^2}$$

$$\mathbb{P}\left[|X-\mathbb{E}X| \geq \xi\right] + \mathbb{P}\left[|X-\mathbb{E}X| < \xi\right] \leq \frac{\mathrm{Var}\,X}{\xi^2}$$

$$\mathbb{P}\left(|X-\mathbb{E}X| \geq \xi\right) \leq \frac{\mathrm{Var}X}{\xi^2}$$

■

CHERNOFF BOUNDS

Given X, the moment generating function is $M_X(t) = \mathbb{E}(e^{tX})$

Thm: Let X be a random variable, then $\forall \xi \in \mathbb{R}$

if $t>0$ $\quad \mathbb{P}(X \geq \xi) \leq e^{-\xi t} \cdot M_X(t)$

if $t<0$ $\quad \mathbb{P}(X \leq \xi) \leq e^{-\xi t} \cdot M_X(t)$

where $M_X(\cdot)$ is the moment generating function

PROOF: Let $t>0$,

$$\mathbb{P}(X>\xi) = \mathbb{P}(e^{tx} > e^{t\xi}) \underset{\text{Markov}}{\leq} \frac{\mathbb{E}(e^{tx})}{e^{t\xi}} = e^{-\xi t} M_X(t)$$

Let $t<0$

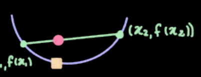
$$\mathbb{P}(X<\xi) = \mathbb{P}(e^{tx} > e^{t\xi}) \leq e^{-\xi t} M_X(t)$$

■

HOEFFDING INEQUALITY

Prop: A function f is convex in $[x_1,x_2]$ if $\forall t \in [0,1]$

$$f(tx_1 + (1-t)x_2) \leq t\cdot f(x_1) + (1-t)\,f(x_2)$$



PROOF: The line between $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is given by

$$r: \frac{y - f(x_1)}{f(x_2) - f(x_1)} = \frac{x - x_1}{x_2 - x_1}$$

$$\downarrow$$

$$y - f(x_1) = \frac{f(x_2)-f(x_1)}{x_2 - x_1}x + \frac{f(x_1)-f(x_2)}{x_2 - x_1}\cdot x_1$$

$$y = \frac{x\cdot f(x_2) - x\cdot f(x_1) + \cancel{x_1 f(x_1)} - x_1\cdot f(x_2) + x_2\cdot f(x_1) - \cancel{x_1 f(x_1)}}{x_2 - x_1}$$

$$y = \frac{x - x_1}{x_2 - x_1}\cdot f(x_2) + \frac{x_2 - x}{x_2 - x_1}\cdot f(x_1) = r(x)$$

Since f is convex, $\forall x \in [x_1, x_2]$:

$$f(x) \leq r(x)$$

$$\downarrow \quad x = tx_1 + (1-t)x_2, \quad \forall t \in [0,1]$$

$$f(tx_1 + (1-t)x_2) \leq \frac{tx_1 + (1-t)x_2 - x_1}{x_2 - x_1}\cdot f(x_2) + \frac{x_2 - tx_1 - (1-t)x_2}{x_2 - x_1}\,f(x_1)$$

$$f(tx_1 + (1-t)x_2) \leq (t-1)\,f(x_2) + t\cdot f(x_1)$$

$$f(tx_1 + (1-t)x_2) \leq t\cdot f(x_1) - (t-1)\,f(x_2)$$

■

Prop: $\forall a,b,\quad 4ab \leq (a+b)^2$

PROOF:

$$0 \leq (a-b)^2$$
$$0 \leq a^2 - 2ab + b^2$$
$$0 \leq a^2 + 2ab + b^2 - 4ab$$
$$4ab \leq a^2 + 2ab + b^2$$
$$4ab \leq (a+b)^2$$

Lemma : Hoeffding

Let be Y r.v. such that $Y \in [a,b]$ a.s., then

$$\mathbb{E}\left[e^{t(Y-\mathbb{E}Y)} \right] \leqslant e^{\frac{t^2(b-a)^2}{8}}$$

PROOF: The function $e^{\lambda x - \mathbb{E}X}$ is a convex function, hence [in] a,b for $x = b - t(b-a)$ we have for the previous proposition:

$$f(x) \leqslant \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$$

$$e^{\lambda(x-\mathbb{E}X)} \leqslant \frac{b-x}{b-a} e^{\lambda a - \lambda \mathbb{E}X} + \frac{x-a}{b-a} e^{\lambda b - \lambda \mathbb{E}X}$$

$$\Big\downarrow \ \mathbb{E} \text{ is monotonicity and linear}$$

$$\mathbb{E}\left(e^{\lambda(X-\mathbb{E}X)}\right) \leqslant \frac{b-\mathbb{E}X}{b-a} \cdot e^{\lambda a - \lambda \mathbb{E}X} + \frac{\mathbb{E}X-a}{b-a} e^{\lambda b - \lambda \mathbb{E}X} =$$

$$= e^{\lambda(a-\mathbb{E}X)} \left[\frac{b-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-a} \cdot e^{\lambda(b-a)} \right] =$$

$$= e^{\lambda(a-\mathbb{E}X) + \ln\left(\frac{b-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-a} e^{\lambda(b-a)}\right)} =$$

$$\Big\downarrow \text{ let } \quad L(h) := \frac{h \cdot (a-\mathbb{E}X)}{b-a} + \ln\left(\frac{b-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-a} e^{h}\right)$$

$$= e^{L(\lambda(b-a))} \leqslant$$

$$L(0) = 0 + \ln(1) = 0$$

$$L'(h) = \frac{a-\mathbb{E}X}{b-a} + \frac{1}{\frac{b-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-a} \cdot e^{h}} \cdot \frac{\mathbb{E}X-a}{b-a} \cdot e^{h} =$$

$$= \frac{a-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-\mathbb{E}X + (\mathbb{E}X-a)e^{h}} \, e^{h} =$$

$$= (a-\mathbb{E}X) \left[\frac{1}{b-a} - \frac{e^{h}}{b-\mathbb{E}X + (\mathbb{E}X-a)e^{h}} \right]$$

$$L'(0) = \frac{a-\mathbb{E}X}{b-a} + \frac{\mathbb{E}X-a}{b-a} = 0$$

$$L''(h) = -(a-\mathbb{E}X) \cdot \frac{e^{h}\left(b-\mathbb{E}X+(\mathbb{E}X-a)e^{h}\right) - e^{h} \cdot (\mathbb{E}X-a)e^{h}}{\left(b-\mathbb{E}X + (\mathbb{E}X-a)e^{h}\right)^2} =$$

$$= -(a-\mathbb{E}X) \cdot e^{h} \frac{b-\mathbb{E}X + (\mathbb{E}X-a)e^{h} - (\mathbb{E}X-a)e^{h}}{\left(b-\mathbb{E}X + (\mathbb{E}X-a)e^{h}\right)^2} =$$

$$= e^{h} \frac{(b-\mathbb{E}X)(\mathbb{E}X-a)}{\left[b-\mathbb{E}X + e^{h}(\mathbb{E}X-a)\right]^2}$$

$$\Big\downarrow \quad B = b-\mathbb{E}X, \quad A = (\mathbb{E}X-a)e^{h}$$

$$= \frac{A \cdot B}{(A+B)^2} \leqslant \frac{A \cdot B}{4 \cdot A \cdot B} = \frac{1}{4}$$

$$(A+B)^2 \geqslant 4AB$$

By Taylor Theorem (with Lagrange remainder), we get:

$$L(h) = L(0) + L'(0) \cdot h + \frac{L''(\xi)}{2} \cdot h^2 \leqslant \frac{1}{8} \cdot h^2$$

$$\leqslant e^{L(\lambda(b-a))} \leqslant e^{\frac{1}{8}\lambda^2 (b-a)^2} \qquad \blacksquare$$

Thm: Let $X_1,\dots,X_n$ i.r.v. such that $\forall i=1,\dots,n$, $a_i \le X_i \le b_i$ almost surely. Let $S_n = X_1+\dots+X_n$, then $\forall \xi > 0$

$$\mathbb{P}(S_n - \mathbb{E}(S_n) > \xi) \le e^{-\frac{2\xi^2}{\sum (b_i-a_i)^2}}$$

PROOF:

$$\mathbb{P}(S_n - \mathbb{E}(S_n) > \xi) \overset{\substack{\text{Chernoff bound} \\ \forall t>0 \\ Y := S_n - \mathbb{E}(S_n)}}{\le} e^{-\xi t} \cdot M_Y(t) =$$

$$= e^{-\xi t} \cdot \mathbb{E}\left(e^{t(S_n - \mathbb{E}S_n)}\right) =$$

$$= e^{-\xi t} \cdot \mathbb{E}\left(e^{\sum t(X_i - \mathbb{E}X_i)}\right) =$$

$\downarrow$ X_i indipendent

$$= e^{-\xi t} \cdot \prod_i \left(\mathbb{E}\left(e^{t(X_i - \mathbb{E}X_i)}\right)\right) \le$$

$\downarrow$ Hoeffding Lemma apply n-times

$$\le e^{-\xi t} \cdot \prod_i \cdot e^{t^2 \frac{(b_i-a_i)^2}{8}} \le$$

$$\le e^{-\xi t + \frac{t^2}{8} \cdot \sum (b_i-a_i)^2}$$

$\downarrow$ let $t := \frac{4\xi}{\sum (b_i-a_i)^2}$

$$\le e^{-\xi \frac{4\xi}{\sum(b_i-a_i)^2} + \frac{16\xi^2}{8} \frac{\sum (b_i-a_i)^2}{(\sum (b_i-a_i)^2)^2}} =$$

$$\le e^{-\frac{4\xi^2}{\sum(b_i-a_i)^2} + \frac{2\xi^2}{\sum(b_i-a_i)^2}} =$$

$$= e^{-\frac{2\xi^2}{\sum (b_i-a_i)^2}}$$

$\blacksquare$

Lemma: Popoviciu inequality

Let be Y r.v. such that $Y \in [a,b]$ a.s., then

$$\mathbb{V}ar(Y) \le \frac{(b-a)^2}{4}$$

PROOF: From the hypothesis, we know $\mathbb{P}(a \le Y \le b) = 1$, then

$$0 \le \mathbb{E}\,[\underbrace{(b-Y)(Y-a)}_{\text{positive r.v.}}] = \mathbb{E}\left[bY - ab - Y^2 + aY\right] =$$

$$\le (b+a)\,\mathbb{E}Y - ab - \mathbb{E}Y^2$$

Hence we get

$$\mathbb{E}Y^2 \le (b+a)\,\mathbb{E}Y - ab$$

$$\mathbb{V}ar(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 \le (b+a)\,\mathbb{E}Y - ab - \mathbb{E}Y^2 \le$$

$$\le (b - \mathbb{E}Y)(\mathbb{E}Y - a) \le$$

$\downarrow$ $xy \le \left(\frac{x+y}{2}\right)^2$

$$\le \frac{(b - \mathbb{E}Y + \mathbb{E}Y - a)^2}{4} = \frac{(b-a)^2}{4}$$

$\blacksquare$

BENNETT'S INEQUALITY

Thm: Let $X_1, \dots, X_n$ i.r.v. such that $\forall i = 1, \dots, n$,

(a) $\mathbb{V}ar\, X_i$ exists finite

(b) $|X_i - \mathbb{E}X_i| \leq M_i$ a.s.

Let $S_n := X_1 + \dots + X_n$, $M = \max\{M_i\}$, then $\forall \xi > 0$

$$\mathbb{P}(S_n - \mathbb{E}S_n > \xi) \leq e^{-\frac{\mathbb{V}ar\, S_n}{M^2} \cdot L\left(\frac{M\xi}{\mathbb{V}ar\, S_n}\right)}$$

where $L(u) := (1+u) \cdot \log(1+u) - u$

PROOF: $\mathbb{P}(S_n - \mathbb{E}S_n > \xi) \overset{\text{Chernoff bounds}}{\leq}$

$$\leq e^{-\xi t}\, \mathbb{E}\left(e^{t(S_n - \mathbb{E}S_n)}\right) =$$

$$= e^{-\xi t} \prod \mathbb{E}\left(e^{t(X_i - \mathbb{E}X_i)}\right) =$$

consider the function $\varphi : u \mapsto \frac{e^u - u - 1}{u^2}$

$$\lim_{u \to 0} \varphi(u) = \lim_{u \to 0} \frac{e^u - 1 - u}{u^2} =$$

de L'hôpital

$$= \lim_{u \to 0} \frac{e^u - 1}{2 \cdot u} = \frac{1}{2}$$

$$\varphi'(u) = \frac{(e^u - 1)u^2 - 2u(e^u - 1 - u)}{u^4} =$$

$$= \frac{u(e^u - 1) - 2(e^u - 1 - u)}{u^3} =$$

$$= \frac{e^u(u-2) + u + 2}{u^3} \underset{\text{from analysis}}{>} 0$$

Hence $\forall i = 1, \dots, n$,

$$t(X_i - \mathbb{E}X_i) \leq tX_i \leq t \cdot M$$

φ is increasing

$$\varphi(t(X_i - \mathbb{E}X_i)) \leq \varphi(tM)$$

$$\frac{e^{t(X_i - \mathbb{E}X_i)} - t(X_i - \mathbb{E}X_i) - 1}{t^2 (X_i - \mathbb{E}X_i)^2} \leq \frac{e^{tM} - tM - 1}{t^2 M^2}$$

$$e^{t(X_i - \mathbb{E}X_i)} \leq \frac{(X_i - \mathbb{E}X_i)^2}{M^2} \cdot \left(e^{tM} - tM - 1\right) + t(X_i - \mathbb{E}X_i) + 1$$

$$\leq e^{-\xi t} \cdot \prod \mathbb{E}\left(\frac{(X_i - \mathbb{E}X_i)^2}{M^2}(e^{tM} - tM - 1) + t(X_i - \mathbb{E}X_i) + 1\right) =$$

$$= e^{-\xi t} \cdot \prod \left[\frac{e^{tM} - tM - 1}{M^2}\, \mathbb{E}(X_i - \mathbb{E}X_i)^2 + t \cdot \left[\mathbb{E}(X_i) - \mathbb{E}X_i\right] + 1\right] =$$

$$= e^{-\xi t} \cdot \prod \left[\frac{e^{tM} - tM - 1}{M^2} \cdot \mathbb{V}ar(X_i) + 1\right]$$

$1 + x \leq e^x$

$$\leq e^{-\xi t} \cdot \prod \left[e^{\frac{\mathbb{V}ar\, X_i}{M^2} \cdot (e^{tM} - tM - 1)}\right] =$$

$$= e^{-\xi t} \cdot e^{\frac{e^{tM} - tM - 1}{M^2} \cdot \mathbb{V}ar\, S_n} =$$

$$= e^{-\xi t + \frac{e^{tM} - tM - 1}{M^2} \mathbb{V}ar\, S_n} =$$

We want to find the optimal t such that the error is minimal

$$\Psi(t) = e^{-\xi t + \frac{e^{tM} - tM - 1}{M^2} \mathbb{V}\mathrm{ar}\, S_n}$$

$$\Psi'(t) = e^{-\xi t + \frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \cdot (e^{tM} - tM - 1)} \cdot$$

$$\cdot \left(-\xi + \frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \left(e^{tM} \cdot M - M \right) \right)$$

A minimum is for

$$-\xi + \frac{\mathbb{V}\mathrm{ar}\, S_n}{M} \left(e^{tM} - 1 \right) = 0$$

$$e^{tM} - 1 = \frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n}$$

$$t = \frac{1}{M} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right)$$

$$\mathbb{P}(S_n - \mathbb{E} S_n > \xi) \leq$$

$$\leq e^{\left\{ -\xi \frac{1}{M} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) + \frac{e^{\cancel{M} \cdot \frac{1}{\cancel{M}} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right)} - \frac{1}{\cancel{M}} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) \cdot \cancel{M} - 1}{M^2} \cdot \mathbb{V}\mathrm{ar}\, S_n \right\}}$$

$$\leq e^{\left\{ -\frac{\xi}{M} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) + \frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \left[\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} \cancel{+1} \cancel{-1} - \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) \right] \right\}}$$

$$= e^{\frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \left\{ -\frac{\xi M}{\mathbb{V}\mathrm{ar}\, S_n} \cdot \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) + \frac{\xi M}{\mathbb{V}\mathrm{ar}\, S_n} - \log\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} + 1 \right) \right\}}$$

$$= e^{-\frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \cdot \left\{ \left(1 + \frac{\xi M}{\mathbb{V}\mathrm{ar}\, S_n} \right) \cdot \log\left(1 + \frac{\xi M}{\mathbb{V}\mathrm{ar}\, S_n} \right) - \frac{\xi M}{\mathbb{V}\mathrm{ar}\, S_n} \right\}}$$

$$\leq e^{-\frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2} \cdot L\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n} \right)}$$

where $L(u) := (1+u) \cdot \log(1+u) - u$ $\blacksquare$

BERNSTEIN'S INEQUALITY

Prop: $\forall u > 0$ $\quad (1+u)\log(1+u) - u \geq \dfrac{u^2}{2\left(1+\frac{u}{3}\right)}$

PROOF: Let $f(x) = \log(1+x)$, the McLaurin expansion is

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{6} + \cdots = \sum_{i=1}^{+\infty} \frac{(-1)^{i+1}}{i!} x^i$$

Then

$$
\begin{aligned}
&(1+u)\cdot\log(1+u) - u = \\
&= (1+u)\left(u - \frac{u^2}{2} + \frac{u^3}{6} + \sum_{i=4}^{+\infty} \frac{(-1)^{i+1}}{i!} u^i\right) - u = \\
&= \cancel{u} - \frac{u^2}{2} + \frac{u^3}{6} + \sum_{i=4}^{+\infty} \frac{(-1)^{i+1}}{i!} u^i + \\
&\quad u^2 - \frac{u^3}{2} + \sum_{i=3}^{+\infty} \frac{(-1)^{i+1} u^{i+1}}{i!} - \cancel{u} = \\
&= \frac{u^2}{2} - \frac{u^3}{3} + \sum_{i=4}^{+\infty} \frac{(-1)^{i+1}}{i!} u^i + \sum_{j=4}^{+\infty} \frac{(-1)^j}{(j-1)!} u^j = \\
&= \frac{u^2}{2} - \frac{u^3}{3} + \sum_{i=4}^{+\infty} \frac{(-1)^i}{i!}(i-1)\cdot u^i = \\
&= \frac{u^2}{2} - \frac{u^3}{3} + o(u^4)
\end{aligned}
$$

$$
\begin{aligned}
\frac{u^2}{2\left(1+\frac{u}{3}\right)} &= \frac{3}{2}u^2\left[\frac{1}{3+u}\right] = \\
&= \frac{3}{2}u^2\left[\frac{1}{3} - \frac{1}{9}\cdot u + o(u^2)\right] = \\
&= \frac{u^2}{2} - \frac{u^3}{6} + o(u^4)
\end{aligned}
$$

$\blacksquare$

Thm: Let $X_1, \dots, X_n$ i.r.v. such that $\forall i = 1, \dots, n$, $a_i \leq X_i \leq b_i$ almost surely. Let $S_n = X_1 + \cdots + X_n$, and $M = \max\{b_i\}$. Then $\forall \xi > 0$

$$\mathbb{P}(S_n - \mathbb{E}(S_n) > \xi) \leq e^{-\frac{\xi^2}{2}\cdot\frac{1}{\sum \mathbb{V}\mathrm{ar}\, X_i + \frac{1}{3}M\cdot\xi}}$$

PROOF: $\mathbb{P}(S_n - \mathbb{E}S_n > \xi) \overset{\text{Bennett's inequality}}{\leq} e^{-\frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2}\cdot L\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n}\right)} \leq$

from the previous proposition, we have

$$
\begin{aligned}
L\left(\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n}\right) &\geq \frac{M^2\xi^2}{(\mathbb{V}\mathrm{ar}\, S_n)^2}\,\frac{1}{2}\,\frac{1}{1+\frac{1}{3}\frac{M\xi}{\mathbb{V}\mathrm{ar}\, S_n}} = \\
&= \frac{M^2\xi^2}{2\cdot(\mathbb{V}\mathrm{ar}\, S_n)^2}\cdot\frac{\mathbb{V}\mathrm{ar}\, S_n}{\mathbb{V}\mathrm{ar}\, S_n + \frac{1}{3}M\xi} = \\
&= \frac{M^2\cdot\xi^2}{2\cdot\mathbb{V}\mathrm{ar}\, S_n}\cdot\frac{1}{\mathbb{V}\mathrm{ar}\, S_n + \frac{1}{3}M\xi}
\end{aligned}
$$

$$
\begin{aligned}
&\leq e^{-\frac{\mathbb{V}\mathrm{ar}\, S_n}{M^2}\cdot\frac{M^2\xi^2}{2\cdot\mathbb{V}\mathrm{ar}\, S_n}\cdot\frac{1}{\mathbb{V}\mathrm{ar}\, S_n + \frac{1}{3}M\xi}} = \\
&\leq e^{-\frac{\xi^2}{2}\cdot\frac{1}{\mathbb{V}\mathrm{ar}\, S_n + \frac{1}{3}M\xi}}
\end{aligned}
$$

$\blacksquare$

CHERNOFF BOUND for binomial

Lemma: If $g(a)=h(a)$ and $g'(t) \leq h'(t)$ $\forall t \geq a$ $\Longrightarrow$ $g(t) \leq h(t)$

PROOF: $h(t) - g(t) = h(t) - h(a) - (g(t) - g(a)) =$
$$= \int_a^t h'(x) - g'(x) \, dx \geq 0 \qquad \blacksquare$$

Lemma: $\forall t \geq 0$, $\forall p \in [0,1]$:
$$p \cdot e^{t(1-p)} + (1-p) \cdot e^{-tp} \leq e^{t^2/8}$$

PROOF: Let $g(t) := \ln(p \cdot e^t + (1-p))$ and $h(t) := \frac{t^2}{8} + tp$

Then :

☆ $g(0) = 0$
$h(0) = 0$

☆ $g'(t) = \dfrac{p \cdot e^t}{p \cdot e^t + (1-p)}$ $\qquad h'(t) = \frac{t}{4} + p$

$g'(0) = p \qquad h'(0) = p$

☆
$$g''(t) = \frac{pe^t(pe^t + (1-p)) - p \cdot e^t \cdot p \cdot e^t}{(p \cdot e^t + (1-p))^2} =$$
$$= \frac{p \cdot e^t (p \cdot e^t + 1 - p - p \cdot e^t)}{(p \cdot e^t + (1-p))^2} =$$
$$= \frac{p \cdot e^t (1-p)}{(p \cdot e^t + (1-p))^2} \underset{ab \leq \left(\frac{a+b}{2}\right)^2}{\leq} \frac{1}{4}$$

$h''(t) = \frac{1}{4}$
$g''(t) \leq h''(t)$

Applying twice the lemma

$$\left.\begin{array}{l} g'(0) = h'(0) \\ g''(t) \leq h''(t) \end{array}\right\} \left.\begin{array}{l} g'(0) = h'(0) \\ g'(t) \leq h'(t) \end{array}\right\} g(t) \leq h(t)$$

then $\qquad g(t) \leq h(t)$
$$\ln(p \cdot e^t + (1-p)) \leq \frac{t^2}{8} + tp$$
$$pe^t + (1-p) \leq e^{t^2/8} \cdot e^{tp}$$
$$pe^{t-tp} + (1-p)e^{-tp} \leq e^{\frac{t^2}{8}} \qquad \blacksquare$$

Thm: Let X be a random variable, $X \sim Bin(n,p)$, $\forall \xi > 0$
$$\mathbb{P}(X - \mathbb{E}X \geq \xi) \leq e^{-\frac{2\xi^2}{n}}$$
$$\mathbb{P}(X - \mathbb{E}X \leq -\xi) \leq e^{-\frac{2\xi^2}{n}}$$

PROOF: Let $X := \sum X_i$, $\quad X_i \sim Ber(p)$

Let $t > 0$:
$$\mathbb{P}(X - \mathbb{E}X \geq \xi) \overset{\text{Chernoff}}{\leq} e^{-\xi t} \, \mathbb{E}(e^{t(X - \mathbb{E}X)}) =$$
$$\leq e^{-\xi t} \cdot \prod \mathbb{E}(e^{t(X_i - \mathbb{E}X_i)}) =$$
$$\leq e^{-\xi t} \cdot \prod (p \cdot e^{t-tp} + (1-p)e^{t}) =$$
$$= e^{-\xi t} \cdot \prod \left(p \cdot e^{t(1-p)} + e^t - p \cdot e^t\right) =$$
$$= e^{-\xi t} \cdot \prod e^{t^2/8} =$$
$$= e^{-\xi t + \frac{nt^2}{8}} = e^{f(t)} \underset{t \min}{=} e^{-\xi^2/n}$$

where $f(t)$ has minimum for
$$f'(t) = -\xi + \frac{n}{4}t = 0 \implies t = \frac{4\xi}{n}$$
$$f\left(\frac{4\xi}{n}\right) = -\xi \cdot \frac{4\xi}{n} + n \cdot \frac{16\xi^2}{8n^2} = -2\,\xi^2/n \qquad \blacksquare$$

STRONGER CHERNOFF BOUND
for binomial

Lem: Let $X \sim Bin(n,p)$, then:

$$M_X(t) = \left(p \cdot e^t + 1 - p\right)^n \leq e^{\mathbb{E}X\,(e^t - 1)}$$

PROOF: $M_X(t) = \mathbb{E}(e^{tX}) \underset{X \sim Bin(n,p)}{=}$

$$= \sum_{i=0}^{n} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \cdot e^{t \cdot i} =$$

$$= \sum_{i=0}^{n} \binom{n}{i} (p \cdot e^t)^i (1-p)^{n-i} =$$

$$\underset{\text{Newton Formulas}}{=} \left(p \cdot e^t + 1 - p\right)^n =$$

$$= \left(1 + p(e^t - 1)\right)^n =$$

$$\underset{1+x<e^x}{=} e^{np(e^t-1)}$$

$\blacksquare$

Thm: Let $X \sim Bin(n,p)$, $\forall \xi > 0$

$$\mathbb{P}(X - \mathbb{E}X \geq \xi \cdot \mathbb{E}X) \leq \left(\frac{e^{\xi}}{(1+\xi)^{1+\xi}}\right)^{\mathbb{E}X}$$

Futhermore, when $0 < \xi < 1$

$$\mathbb{P}(X - \mathbb{E}X \leq -\xi \cdot \mathbb{E}X) \leq \left(\frac{e^{-\xi}}{(1-\xi)^{1-\xi}}\right)^{\mathbb{E}X}$$

PROOF: Let $\xi > 0$,

$$\mathbb{P}(X - \mathbb{E}X \geq \xi \cdot \mathbb{E}X) =$$

$$= \mathbb{P}(X \geq \mathbb{E}X(1+\xi)) =$$

$$\underset{\text{Chebyshev}}{\leq} e^{-t \cdot (1+\xi) \cdot \mathbb{E}X} \cdot \mathbb{E}(e^{tX}) \leq$$

$$\underset{\text{Lemma}}{\leq} e^{-t \cdot (1+\xi) \cdot \mathbb{E}X} \cdot e^{np(e^t-1)} =$$

$$= \left(e^{-t \cdot (1+\xi) + e^t - 1}\right)^{\mathbb{E}X} \leq$$

$$f(t) = -t(1+\xi) + e^t - 1$$
$$f'(t) = -(1+\xi) + e^t$$

$f'(t) > 0 \quad e^t > 1+\xi \quad t > \ln(1+\xi)$

f has a minimo in $t = \ln(1+\xi)$

$$\leq \left(e^{-\ln(1+\xi)\cdot(1+\xi) + e^{\ln(1+\xi)} - 1}\right)^{\mathbb{E}X} \leq$$

$$\leq \left((1+\xi)^{-(1+\xi)} \cdot e^{(1+\xi)-1}\right)^{\mathbb{E}X} =$$

$$= \left(\frac{e^{\xi}}{(1+\xi)^{1+\xi}}\right)^{\mathbb{E}X}$$

Let $0 < \xi < 1$,

$$\mathbb{P}(X - \mathbb{E}X \leq -\xi \cdot \mathbb{E}X) =$$

$$= \mathbb{P}(X \leq (1-\xi) \cdot \mathbb{E}X) \leq$$

$$\underset{\text{Chernoff bound} \quad t<0}{\leq} e^{-(1-\xi)\cdot \mathbb{E}X \cdot t} \cdot \mathbb{E}(e^{tX}) \leq$$

Lemma

$$\leq e^{-(1-\delta)\cdot \mathbb{E}X \cdot t} \cdot e^{(e^t-1)\cdot \mathbb{E}X} =$$

$$= \left(e^{-(1-\delta)\cdot t + e^t - 1} \right)^{\mathbb{E}X} \leq$$

$$f(t) = -t(1-\delta) + e^t - 1$$

$$f'(t) = -(1-\delta) + e^t$$

$$f'(t) > 0 \qquad e^t > 1-\delta \qquad t > \ln(1-\delta)$$

f has a minimo in $t = \ln(1-\delta)$

$$\leq \left(e^{-(1-\delta)\cdot \log(1-\delta) + e^{\ln(1-\delta)} - 1} \right)^{\mathbb{E}X} =$$

$$= \left((1-\delta)^{-(1-\delta)} \cdot e^{1-\delta-1} \right)^{\mathbb{E}X} =$$

$$= \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^{\mathbb{E}X} \qquad \blacksquare$$