Sub-exponential r.v.'s

Note Title 2015-09-18

What if the r.v.'s have heavier tails?

Example: (Noisy signal)

$$\boxed{y} = \boxed{x} + \boxed{z} \in \mathbb{R}^n$$

observed, signal, noise $z_i \overset{iid}{\sim} N(0,1)$

Error: $\|y-x\|_{\ell_2}^2 = \underbrace{\sum_{i=1}^n z_i^2}$

how does this behave?

z_i^2 is <u>not</u> sub-Gaussian:

$$P(z_i^2 > t) = P(|z_i| > \sqrt{t}) \lesssim \exp\left(\frac{-t}{2}\right)$$

sub-exponential tail

1.4 Sub-exponential r.v.'s

Properties of sub-exp. r.v. X:

① Tails: $P(|X| > t) \le \exp(1 - t/K_1)$ $\quad t \ge 0$

② Moments: $(\mathbb{E}|X|^p)^{1/p} \le K_2 \cdot p$ $\quad p \ge 1$

③ Tails': $\mathbb{E} \exp\left(\frac{|X|}{K_3}\right) \le e$

Note: K_1, K_2, K_3 depend on X.

Lemma For a r.v. X, properties 1 - 3 are equivalent, i.e., $\exists$ an abs. const. C such that, for any X, if prop. i holds w/ K_i then prop. j holds w/ $K_j \le C K_i$ $\forall i,j \in \{1,2,3\}$.

HW: Prove the lemma

Def (Sub-exponential r.v.) A r.v. X which satisfies one of the equiv. properties 1 - 3 is called a sub-exponential r.v. The sub-exponential norm is the smallest K_2 in property 2:

$$\|X\|_{\psi_1} := \sup_{p \ge 1} \frac{(\mathbb{E}|X|^p)^{\frac{1}{p}}}{p}.$$

Examples

1. Exponential r.v.: $X \sim Exp(\lambda)$

$$P(X > t) = e^{-\lambda t} \quad (t \ge 0, \quad x \ge 0)$$

$$\|X\|_{\psi_1} \sim \frac{1}{\lambda} \quad \text{(proof?)}$$

2. The square of any sub-Gauss. r.v. is subexponential:

Lemma X is sub-Gaussian iff X^2 is sub-exponential, and

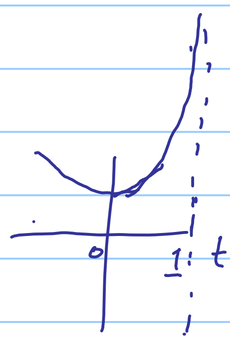
$$\|X\|_{\psi_2}^2 \le \|X^2\|_{\psi_1} \le 2\|X\|_{\psi_2}^2$$

HW Prove this.

To prove a dev. ineq. we wish to control MGF. However MGF may be unbounded!

Example: $Z \sim \mathrm{Exp}(1)$, $X := Z - \mathbb{E}Z$

MGF: $$\mathbb{E}\exp(tX) = \begin{cases} \dfrac{e^{-t}}{1-t}, & t<1 \\ \text{not defined}, & t>1 \end{cases}$$



Can we control MGF in some interval?

Lemma (MGF of sub-exponential)

Let X be a sub-exp. r.v. w/ $\mathbb{E}X=0$. Then

$$\mathbb{E}\exp(tX) \le \exp\left(C t^2 \|X\|_{\psi_1}^2\right) \quad \text{for } |t| \le \frac{c}{\|X\|_{\psi_1}}$$

Proof: WLOG, assume $\|X\|_{\psi_1} = 1$

Taylor series:

$$\mathbb{E}\exp(tX) = \mathbb{E}\left[1 + tX + \sum_{p=2}^{\infty} \frac{(tX)^p}{p!}\right]$$

$$= 1 + \cancel{t\mathbb{E}X}^{\,0} + \sum_{p=2}^{\infty} \frac{t^p \mathbb{E}X^p}{p!}$$

$$\leq 1 + \sum_{p=2}^{\infty} \frac{t^p p^p}{\left(\frac{p}{e}\right)^p} \qquad \left(p! \geq \left(\frac{p}{e}\right)^p \text{ by stirling } \neq\right)$$

$$= 1 + \sum_{p=2}^{\infty} (et)^p$$

If $|et| \leq \frac{1}{2}$, the geom. series converges, and we obtain

$$\mathbb{E}\exp(tX) \leq 1 + 2(et)^2 \leq \exp(2e^2t^2)$$

QED

MGF $\Rightarrow$ dev. $\neq$

$\boxed{\text{Thm (Bernstein's } \neq \text{ for sub-exp. r.v.'s)}}$

Let $X_1, \dots X_n$ be indep., mean-zero, sub-exp. r.v.'s. Then $\forall \{a_1, \dots a_n\} \in \mathbb{R}^n$,

$$P\left(\left|\sum_{i=1}^{n} a_i X_i\right| > t\right) \leq 2\exp\left[-C\min\left(\frac{t^2}{K^2\|a\|_2^2}, \frac{t}{K\|a\|_\infty}\right)\right], \quad t \geq 0$$

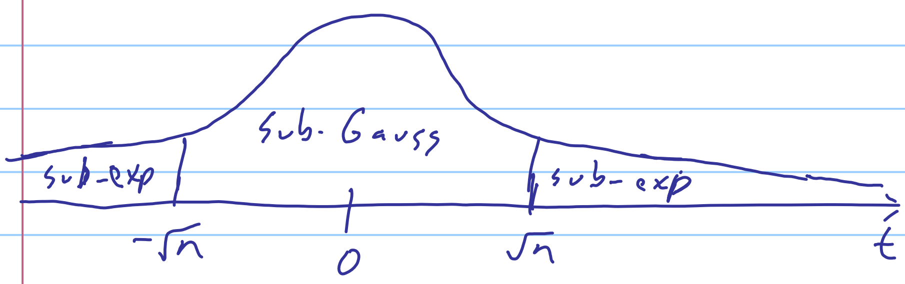
where $K = \max_i \|X_i\|_{\psi_1}$

Roles of the two tails:

Special case: $k=1$, $a_i = \frac{1}{\sqrt{n}}$ $\Rightarrow \|a\|_2 = 1$, $\|a\|_\infty = \frac{1}{\sqrt{n}}$

$$P\left(\left|\frac{1}{\sqrt{n}}\sum_{i=1} X_i\right| > t\right) \leq 2\exp\left(-c\min\left(t^2, t\sqrt{n}\right)\right)$$

$$= 2\begin{cases} e^{-ct^2}, & t \leq \sqrt{n} \quad \text{sub-Gauss. tail} \\ e^{-c\sqrt{n}\cdot t}, & t \geq \sqrt{n} \quad \text{sub exp. tail} \end{cases}$$



Reason for sub-exp tail? Caused by fluctuation of single r.v.:

$$P\left(\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right| > t\right) \geq \frac{1}{2} P\left(|X_i| > \sqrt{n}\, t\right) \quad \left(\text{for } X_i \text{ symm.}\right)$$

$$\sim \exp\left(-t\sqrt{n}\right)$$

- Sub-exp part of tail: no interaction between X_i's.

- Sub-Gauss. part of tail: lots of interaction

<u>Proof of thm:</u> WLOG, $K=1$. Set $S := \sum_{i=1}^n a_i X_i$

Start as in proof of Hoeffding ≠

$$P(S > t) = P\left(e^{\lambda S} > e^{\lambda t}\right) \quad (\lambda \text{ will be determined later})$$

$$\le e^{-\lambda t}\, \mathbb{E}\, e^{\lambda S}$$

$$= e^{-\lambda t} \prod_{i=1}^n \mathbb{E}\, e^{\lambda a_i X_i} \qquad (*)$$

To use the lemma bounding MGF, choose λ small enough that $\lambda \|a\|_\infty \le c \Rightarrow$ for each i,

$\lambda |a_i| \le c \le \frac{c}{\|X_i\|_{\psi_1}}$. Thus, by the lemma, $\mathbb{E}\, e^{\lambda a_i X_i} \le \exp(C \cdot \lambda^2 a_i^2)$

$$\Rightarrow (*) \le \exp\left(-\lambda t + C\lambda^2 \sum_{i=1}^n a_i^2\right)$$

Minimize w/ respect to λ under constraint $\lambda \le \frac{c}{\|a\|_\infty}$.

$$\underset{P(S>t)}{\Rightarrow (*)} \le \exp\left[-c \min\left(\frac{t^2}{\|a\|_2^2}, \frac{t}{\|a\|_\infty}\right)\right] = (**)$$

Control $P(-S > t)$ w/ same steps.

$$P(|S| > t) \le P(S > t) + P(-S > t) \le 2(**)$$

QED