Lecture 2: Subgaussian RVs

Isabella Zhu

10 February 2025

§1 Motivation

CLT tells us that if $X_1, \dots X_n \overset{\text{i.i.d.}}{\sim} P$ then

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \sim N(0,1)$$

as $n \to \infty$.

However, if $P = N(\mu, \sigma^2)$, then we can say that

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \sim N(0,1)$$

for **all** n.

So we we meet in the middle. Intuitively, subgaussian distributions satisfy

$$\text{Gaussian distributions} \subseteq \text{Subgaussian distributions} \subseteq \text{All distributions}$$

Morally, subgaussian distributions *look like Gaussians*.

§2 Definitions

§2.1 Gaussian Tails

Let $z \sim N(0,1)$, we know that $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ (density). We care about *tail bounds*, i.e.

$$\mathbb{P}(z > t) = \int_t^\infty \phi(z) dz$$

which has no closed form.

To upper bound this, we can do

$$\begin{aligned}
\mathbb{P}(z > t) &= \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&\le \int_t^\infty \frac{z}{t} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \frac{e^{-t^2/2}}{t\sqrt{2\pi}}
\end{aligned}$$

It follows that

Corollary 2.1

(Mill's Ratio Inequality) We have $\mathbb{P}(|z| > t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}$.

It follows that CLT states that

$$\lim_{n\to\infty} \mathbb{P}(|\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}| > t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}$$

How do we generalize this for n not converging to ∞?

Theorem 2.2

(Berry-Esseen) If $\mathbb{E}[|X|^3] < \infty$, then we have

$$|\mathbb{P}(|\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}| > t) - \Phi(t)| \leq \frac{c}{\sqrt{n}}$$

for all n, where c is some constant.

Then, we can bound using triangle inequality:

$$\mathbb{P}(|\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}| > t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t} + \frac{c}{\sqrt{n}}.$$

§2.2 Moment Generating Functions (MGFs)

Definition 2.3. The **MGF** $M_X(.)$ of random variable X is the function

$$s \to M_X(s) = \mathbb{E}[e^{sX}]$$

for all $s \in \mathbb{R}$. For $s \in \mathbb{R}^d$, we have

$$s \to M_X(s) = \mathbb{E}[e^{\langle s, X\rangle}]$$

Proposition 2.4

The MFG of $Z \sim N(0,1)$ is $M_Z(s) = e^{s^2/2}$.

§2.3 Subgaussian Random Variables

Definition 2.5. A random variable X is subGaussian with variance proxy σ^2 if

1. $\mathbb{E}[X] = 0$
2. $\mathbb{E}[e^{sX}] \leq e^{\sigma^2 s^2/2}$

We write this as $X \sim \text{subG}(\sigma^2)$. Note that this is an abuse of notation, as subG refers to a *family* of distributions, not just one distribution.

Remark 2.6. It is not actually necessary for subGaussians to be centered. We are just doing that for convenience in this class.

Remark 2.7. The following are trivially true:

- $X \sim \text{subG}(\sigma^2)$ implies $-X \sim \text{subG}(\sigma^2)$.
- $X \sim \text{subG}(1)$ implies $\sigma X \sim \text{subG}(\sigma^2)$.

Proposition 2.8

Let X be s.t. $\mathbb{E}[X] = 0$ and $var(X) = 1$. The following are equivalent:

1. $\mathbb{E}[e^{sX}] \leq e^{c_1 s^2}$.
2. $\mathbb{P}(|X| > t) \leq 2e^{-c_2 t^2}$.
3. $||X||_p = (\mathbb{E}[|X|^p])^{\frac{1}{p}} \leq c_3\sqrt{p}$ for all $p \in \mathbb{N}$.
4. $\mathbb{E}[e^{sX^2}] \leq c_4$, $\forall s \in (0, c_i)$.

We will show $1 \to 2$ below. On pset, show $2 \to 3$, $3 \to 1$, and some chain of equivalences involving 4.

For $t > 0$, we have

$$\begin{aligned}
\mathbb{P}(X > t) &= \min_s \mathbb{P}(e^{sX} > e^{st}) \\
&\leq \min_s \mathbb{E}[e^{sX}]e^{-st} \\
&\leq \min_s e^{c_2 s^2 - st} \\
&= e^{-c_2 t^2/2}
\end{aligned}$$

The trick here is that we can choose the optimal s for our purposes. Symmetry for $\mathbb{P}(X < t)$ to get final bound of $2e^{-c_2 t^2/2}$ as desired.

§2.4 Examples of Subgaussians

Example 2.9

$N(0, \sigma^2)$. Duh.

Example 2.10

$|X| \leq n$ and $\mathbb{E}[X] = 0$.

Then the question becomes what is the max amount of variance a r.v. in a finite interval can have? For this we will use Hoeffding's Inequality.

Example 2.11

$X \sim \text{Rad}(\frac{1}{2}) = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases}$

We can check (4) directly: $\mathbb{E}[e^{sX^2}] = e^S$.

We can also check (1) directly via Taylor expansion

$$\mathbb{E}[e^{sX}] = \frac{e^s + e^{-s}}{2} = \frac{1}{2}\sum_{k=0}^{\infty}\frac{s^{2k}}{(2k)!} \le \sum_{k=0}^{\infty}\frac{s^{2k}}{k!2^k} = e^{s^2/2}$$

§3 Tail Bound Results

Theorem 3.1

(Hoeffding's Lemma). Let X be such that $\mathbb{E}[X] = 0$ and $X \in [a, b]$. Then, we have

$$\mathbb{E}[e^{sX}] \le e^{s^2(b-a)^2/8},$$

for every $s \in \mathbb{R}$. Equivalent to $X \sim \text{subG}(\frac{(b-a)^2}{4})$.

Proof. We will define log MGF to be

$$\Psi(s) = \log \mathbb{E}[e^{sX}].$$

We will show that $\Psi(s) \le cs^2$ for some c and wishfully pray that $c = (b-a)^2/8$.

We will show that $\Psi''(s)$ is bounded by a constant and then integrate twice.

We know that $\Psi(0) = 0$ and

$$\Psi'(s) = \frac{\mathbb{E}[Xe^{sX}]}{\mathbb{E}[e^{sX}]}$$

which implies that $\Psi'(0) = \mathbb{E}[\mathbb{X}] = 0$.

We can now compute the second derivative

$$\Psi''(s) = \frac{\mathbb{E}[X^2e^{sX}]}{\mathbb{E}[e^{sX}]} - (\mathbb{E}[\frac{Xe^{sX}}{\mathbb{E}[e^{sX}]}])^2 = var(Y)$$

where we define a new random variable Y such that Y has density $\frac{e^{sX}}{\mathbb{E}[e^{sX}]}$ w.r.t. P_X.

We can conclude that $Y \in [a, b]$ almost surely. We have

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

Then we can use FToC to get $\Psi(s)$ and it's pretty clear that we are done.