Sub-Gaussian $\mathbb{E}\, e^{-\alpha(W-\mu)} \le e^{\alpha^2\sigma^2/2}$

$\to \mathbb{P}(W-\mu \ge t) \le e^{-t^2/(2\sigma^2)}$

Linear combinations of sub-G ~~var~~ r.v.'s are sub-G.

<u>Prop 12</u> Let $(W_i)_{i=1}^n$ be a sequence of independent mean-zero sub-G r.v.-s with respective parameters $(\sigma_i)_{i=1}^n$ and let $\gamma \in \mathbb{R}^n$. Then $\gamma^T W$ is sub-G with parameter $\left(\sum_{i=1}^n \sigma_i^2 \gamma_i^2\right)^{1/2}$.

<u>Proof</u>
$$\mathbb{E} \exp\left(\alpha \sum_{i=1}^n \gamma_i W_i\right) = \mathbb{E} \prod_{i=1}^n \exp(\alpha\gamma_i W_i)$$
$$= \prod_{i=1}^n \mathbb{E} \exp(\alpha \gamma_i W_i)$$
$$\le \prod_{i=1}^n \exp(\alpha^2 \gamma_i^2 \sigma_i^2/2)$$
$$= \exp\left(\alpha^2 \sum_{i=1}^n \gamma_i^2\sigma_i^2/2\right) \quad \square$$

General version of the probability bd required for Thm 9.

<u>Lemma 13</u> Suppose $(\varepsilon_i)_{i=1}^n$ are independent mean-zero sub-G r.v.-s with sub-G parameters $\sigma > 0$. (Note this includes $\varepsilon \sim N_n(0, \sigma^2 I)$). Let $\lambda = A\sigma\sqrt{\frac{\log p}{n}}$. Then
$$\mathbb{P}(\|X^T\varepsilon\|_\infty / n \le \lambda) \ge 1 - 2p^{-(A^2/2)-1})$$

<u>Proof</u> $\mathbb{P}(\|X^T\varepsilon\|_\infty/n > \lambda) \le \sum_{j=1}^p \mathbb{P}(|x_j^T\varepsilon|/n > \lambda) \qquad (*)$

But $\pm x_j^T \varepsilon / n$ are sub-G with parameter $\left(\sigma^2 \sum_{i=1}^n \left(\frac{X_{ij}}{n}\right)^2\right)^{1/2} = \sigma/\sqrt{n}$.

Thus by prop 10, $(*) \le 2p\, e^{-\lambda^2 n/(2\sigma^2)}$
$$= 2p \exp(-A^2 \log p/2) = 2p^{1-A^2/2} \quad \square$$

The following will be helpful for later results

<u>Defn 4</u> (Bernstein's condition)
We say a r.v. W satisfies Bernstein's condition with parameter (σ, b) where $\sigma, b > 0$ if
$$\mathbb{E}|W - \mathbb{E}W|^k \le \tfrac{1}{2} k!\, \sigma^2 b^{k-2} \quad \text{for } k = 2, 3, \ldots$$

<u>Prop 14</u> (Bernstein's inequality) Let $W_1, \ldots, W_n$ be independent r.v.-s with $\mathbb{E}W_i = \mu$. Suppose each W_i satisfies Bernstein's condition with parameter (σ_i, b).

Then
$$\mathbb{E}\, e^{\alpha(W_i-\mu)} \le \exp\left(\frac{\alpha^2\sigma^2/2}{1-b|\alpha|}\right) \quad \text{for all } |\alpha| < \frac{1}{b}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n W_i - \mu \ge t\right) \le \exp\left(-\frac{nt^2}{2(\sigma^2+bt)}\right) \quad \forall\, t \ge 0$$

Proof Fix i and let $W = W_i$ and let $|\alpha| < \frac{1}{b}$.

$$\mathbb{E}\, e^{\alpha(W-\mu)} = 1 + \alpha\,\mathbb{E}(W-\mu) + \sum_{k=2}^\infty \frac{\alpha^k\,\mathbb{E}(W-\mu)^k}{k!} \quad \text{(by dominated convergence)}$$

$$\le 1 + \sum_{k=2}^\infty |\alpha|^k \frac{\mathbb{E}|W-\mu|^k}{k!}$$

$$\le 1 + \frac{\sigma^2\alpha^2}{2}\sum_{k=2}^\infty |\alpha|^{k-2} b^{k-2} = 1 + \frac{\sigma^2\alpha^2}{2}\frac{1}{1-|\alpha|b}$$

$$\overset{1+u\le e^u}{\le} \exp\left(\frac{\sigma^2\alpha^2/2}{1-|\alpha|b}\right)$$

For the probability bd

$$\mathbb{E}\exp\left(\sum_{i=1}^n \alpha(W_i-\mu)/n\right) = \prod_{i=1}^n \mathbb{E}\exp(\alpha(W_i-\mu)/n)$$

$$\le \exp\left(n\,\frac{(\alpha/n)^2\sigma^2/2}{1-b|\alpha/n|}\right) \qquad \left|\frac{\alpha}{n}\right| < \frac{1}{b}$$

Chernoff bd

$$\mathbb{P}\left(\frac{1}{n}\sum W_i - \mu \ge t\right) \le e^{-\alpha t}\exp\left(n\,\frac{(\alpha/n)^2\sigma^2/2}{1-b|\alpha/n|}\right) \qquad \left|\frac{\alpha}{n}\right| < \frac{1}{b}$$

Setting $\frac{\alpha}{n} = \frac{t}{bt+\sigma^2} \in \left[0, \frac{1}{b}\right)$ gives the result. □

Lemma 15 Let W, Z be mean-zero sub-G r.v.'s with parameters σ_W and σ_Z resp.
Then WZ satisfies Bernstein's condition with param $(8\sigma_W\sigma_Z, 4\sigma_W\sigma_Z)$.

Proof Note that $W^{2k} = \int_0^\infty \mathbb{1}_{\{x < W^{2k}\}}\,dx$

$$\mathbb{E}\,W^{2k} = \int_0^\infty \mathbb{P}(W^{2k} > x)\,dx \qquad \text{subst } t^{2k} = x \qquad 2k\,t^{2k-1}dt = dx$$

$$= 2k\int_0^\infty t^{2k-1}\,\mathbb{P}(|W| > t)\,dt$$

$$\le 4k\int_0^\infty t^{2k-1}\exp\left(-\frac{t^2}{2\sigma_W^2}\right)dt \qquad \text{subst } x = \frac{t^2}{2\sigma_W^2}, \quad \sigma_W^2\,dx = t\,dt$$

$$= 4k\sigma_W^2 \int_0^\infty (2\sigma_W^2 x)^{k-1} e^{-x}\,dx = 2^{k+1}\sigma_W^{2k} k \underbrace{\int_0^\infty x^{k-1} e^{-x}\,dx}_{\Gamma(k) = (k-1)!}$$

$$= 2^{k+1}\sigma_W^{2k}\, k!$$

For any r.v. Y, $\mathbb{E}|Y - \mathbb{E}Y|^k = 2^k\, \mathbb{E}\left|\tfrac{1}{2}Y - \tfrac{1}{2}\mathbb{E}Y\right|^k$

$\leq 2^{k-1}\left(\mathbb{E}|Y|^k + |\mathbb{E}Y|^k\right)$ by Jensen's inequality applied to $t \mapsto |t|^k$

$\leq 2^k\, \mathbb{E}|Y|^k$

Thus $\mathbb{E}|WZ - \mathbb{E}WZ|^k \leq 2^k\, \mathbb{E}|WZ|^k \leq 2^k (\mathbb{E}W^{2k})^{1/2} (\mathbb{E}Z^{2k})^{1/2}$ by C-S

$$\leq 2^{2k+1}\sigma_W^k \sigma_Z^k\, k! = \frac{k!}{2}(8\sigma_W\sigma_Z)^2 (4\sigma_W\sigma_Z)^{k-2} \quad \square$$