Concentration of Measure.

§5.3 The Cramér-Chernoff Method and Subgaussian Random Variables.

Definition 5.2 [Subgaussianity]

A random variable X is σ-subgaussian if for $\forall \lambda \in \mathbb{R}$, it holds that $M_X(\lambda) \leq \exp(\frac{\sigma^2}{2} \cdot \lambda^2)$, where $M_X(\lambda) = \mathbb{E}[\exp(\lambda X)]$ is the moment-generating function (MGF) of X.

Theorem 5.3 [Concentration of Subgaussian Random Variables]

If X is σ-subgaussian, then for any $\varepsilon > 0$,

$$\mathbb{P}(X \geq \varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right).$$

Proof. We take a generic approach called the Cramér-Chernoff method. Let $\lambda > 0$ be some constant to be choosen latter. Then

$$\begin{aligned}
\mathbb{P}(X \geq \varepsilon) &= \mathbb{P}(\exp(\lambda X) \geq \exp(\lambda \varepsilon)) && \text{(Exponentiation)} \\
&\leq \mathbb{E}[\exp(\lambda X)] / \exp(\lambda \varepsilon) && \text{(Markov's ineq.)} \\
&\leq \exp(\tfrac{\sigma^2}{2} \cdot \lambda^2) / \exp(\lambda \varepsilon). && \text{(Def. of subgaussianity)}
\end{aligned}$$

Optimizing over λ (choosing $\lambda = \frac{\varepsilon}{\sigma^2}$ concludes the proof. ∎

Remark. An equivalent and also more convenient form is

$$\mathbb{P}\left(X > \sqrt{2\sigma^2 \cdot \log(1/\delta)}\right) \le \delta$$

holds $\forall \delta \in (0,1)$.

Lemma 5.4 Suppose that X is σ-subgaussian and X_1 and X_2 are independent and σ_1 and σ_2-subgaussian, respectively, then

(a) $\mathbb{E}[X] = 0$ and $\mathbb{V}[X] \le \sigma^2$.

(b) $c \cdot X$ is $|c| \cdot \sigma$-subgaussian for all $c \in \mathbb{R}$.

(c) $X_1 + X_2$ is $\sqrt{\sigma_1^2 + \sigma_2^2}$-subgaussian.

Corollary 5.5 [Concentration of Sum of Subgaussian Random Variables]

Assume that $X_i - \mu$ are independent, σ-subgaussian random variables. Then for $\forall \varepsilon > 0$

$$\mathbb{P}(\hat{\mu} - \mu \geq \varepsilon) \leq \exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right),$$

where $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Remark. A more convenient form is

$$\mathbb{P}\left(\hat{\mu} - \mu \geq \sqrt{\frac{2\sigma^2 \ln \delta^{-1}}{n}}\right) \leq \delta.$$

Remark.

(a) For r.v. X that is not zero-mean, we abuse the notation by saying that X is σ-subgaussian if $X - \mathbb{E}[X]$ is σ-subgaussian.

(b). If $X \in [a,b]$ a.s., then X is $(b-a)/2$-subgaussian.