67939 - Topics in Learning Theory **(Due: 16/06/24)**

Exercise 1

Instructor: Prof. Amit Daniely *Name:* Hadar Tal

Exercise 1

The moment generating function (MGF) of a random variable X is $M_X(\lambda) = \mathbb{E}[e^{\lambda X}]$. Assume that M_X is defined for any λ in a non-empty segment $(-a, a)$. Show that

1. $M_X^{(k)}(0) = \mathbb{E}[X^k]$
2. Show that for a centered Gaussian X with variance σ^2, $M_X(\lambda) = e^{\frac{\lambda^2\sigma^2}{2}}$. In other words, being σ-SubGaussian is equivalent to having MGF that is bounded by the MGF of a centered Gaussian with variance σ^2.
3. Show that if X is uniform over $[a, b]$ then $M_X(\lambda) = \frac{e^{\lambda b} - e^{\lambda a}}{\lambda(b-a)}$.

Exercise 2

1. Show that if X_i is σ_i-SubGaussian for $i = 1, 2$ then $X_1 + X_2$ is $(\sigma_1 + \sigma_2)$-SubGaussian[1].
2. For a sub-Gaussian random variable X, define $\|X\|_{vp}$ as the minimal σ for which X is σ-SubGaussian. Show that $\|\cdot\|_{vp}$ is a norm on the space of centered sub-Gaussian random variables. This norm is called the Proxy Variance norm and $\|X\|_{vp}$ is called the optimal proxy variance of X.

Exercise 3

1. Let X be a σ-SubGaussian random variable. Show that $2\sigma \geq \sqrt{\text{var}(X)}$.
2. If $\|X\|_{vp} = \sqrt{\text{var}(X)}$, then X is called strictly sub-Gaussian. Show that if X is uniform on $\{-1, 1\}$, then it is strictly sub-Gaussian. Conclude that the bound in Hoeffding's lemma is optimal.
3. Show that a linear combination of independent strictly sub-Gaussians is strictly sub-Gaussian.
4. Show that for any $M \geq 1$, there is a random variable X with $\text{var}(X) = 1$ and $\|X\|_{vp} = M$.

Exercise 4

Show that there is a universal constant $C > 0$ for which the following holds. If X is a random variable such that for any $t \geq 0$,

$$\Pr(X - \mathbb{E}[X] \geq t) \leq e^{-\frac{t^2}{2\sigma^2}} \quad \text{and} \quad \Pr(X - \mathbb{E}[X] \leq -t) \leq e^{-\frac{t^2}{2\sigma^2}}$$

then X is $(C\sigma)$-SubGaussian[2].

[1] Use the Hölder inequality ($\mathbb{E}[XY] \leq (\mathbb{E}[X^p])^{1/p}(\mathbb{E}[Y^q])^{1/q}$ if $\frac{1}{p} + \frac{1}{q} = 1$ and $p, q \geq 0$) on $\mathbb{E}[e^{\lambda(X-\mathbb{E}[X])}e^{\lambda(Y-\mathbb{E}[Y])}]$

[2] Hint: You may use the fact that for a non-negative random variable Y, $\mathbb{E}[Y] = \int_0^\infty \Pr(Y \geq x)dx$