Homework Assignment 2

Due Feb 28th midnight

SDS 384-11 Theoretical Statistics

1. Consider a r.v. X such that for all $\lambda \in \mathbb{R}$

$$E[e^{\lambda X}] \leq e^{\frac{\lambda^2 \sigma^2}{2} + \lambda \mu} \tag{1}$$

Prove that:

(a) $E[X] = \mu$.

(b) $\text{var}(X) \leq \sigma^2$.

(c) If the smallest value of σ satisfying the above equation is chosen, is it true that $\text{var}(X) = \sigma^2$? Prove or give a counter-example.

2. Given a symmetric positive semidefinite matrix $Q \in \mathbb{R}^{n \times n}$, consider $Z = \sum_{i,j} Q_{ij} X_i X_j$. When $X_i \sim N(0,1)$, prove the Hanson-Wright inequality.

$$P\left(Z \geq \text{trace}(Q) + t\right) \leq \exp\left(-\min\left\{c_1 t/\|Q\|_{op}, c_2 t^2/\|Q\|_F^2\right\}\right),$$

where $\|Q\|_{op}$ and $\|Q\|_F$ denote the operator and frobenius norms respectively. *Useful facts: Let $\lambda_1 \geq \lambda_2 \geq \dots$ denote the eigenvalues of Q. Remember that $\|Q\|_{op} = \sup_{v:\|v\|=1} \|Qv\| = \lambda_1$. For a PSD matrix Q, trace$(Q) = \sum_i \lambda_i$, and $\|Q\|_F^2 = \sum_i \lambda_i^2$. Hint: The rotation-invariance of the Gaussian distribution and sub-exponential nature of χ^2-variables could be useful.*

3. We will prove properties of subgaussian random variables here. Prove that:

(a) Moments of a mean zero subgaussian r.v. X with variance proxy σ^2 satisfy:

$$E[|X^k|] \leq k 2^{k/2} \sigma^k \Gamma(k/2), \tag{2}$$

where Γ is the gamma function.

(b) If X is a mean 0 subgaussian r.v. with variance proxy σ^2, prove that, $X^2 - E[X^2]$ is a subexponential $(c_1\sigma^2, c_2\sigma^2)$ (we are using the (ν, b) parametrization of subexponentials we did in class, so ν^2 is the variance proxy). Here c_1, c_2 are positive constants.

(c) Consider two independent mean zero subgaussian r.v.s X_1 and X_2 with variance proxies σ_1^2 and σ_2^2 respectively. Show that $X_1 X_2$ is a subexponential r.v. with parameters $(d_1\sigma_1\sigma_2, d_2\sigma_1\sigma_2)$. Here d_1, d_2 are positive constants.

4. Subgaussian and subexponential random variables have moments that are growing suitably so that we can have a bound on the MGF. Consider scalar random variables $X_1, \ldots, X_n$ that are IID samples from some distribution with mean μ. What if all we have is an upper bound on the variance, i.e. $E[(X_1 - \mu)^2] \le \sigma^2 < \infty$ - are there estimators for which we can obtain exponential tail bounds? This is what we will learn through this exercise. Assume $n = mk$ for some positive integers m, k. Divide the data into k disjoint chunks. For each chunk, compute the mean, call this m_i, $i = 1, \ldots, m$. Let your estimator be $\widehat{\mu}_n := \text{median}(\{m_i\}_{i=1}^m)$. We will show that, for some appropriately picked $k = k_\delta$,

$$P\left(|\widehat{\mu}_n - \mu| \ge c\sigma\sqrt{\frac{\log(1/\delta)}{n}}\right) \le \delta \tag{3}$$

where c is a constant.

(a) First show that, for $i \in \{1, \ldots, m\}$ $P\left(|m_i - \mu| \ge \frac{\sigma}{2\sqrt{m}}\right) \le 1/4$

(b) Now find a suitable k as a function of δ, such that Eq 3 holds. *Hint: Use the definition of a median to frame Eq 3 as a failure probability of a sum of k independent Bernoulli(p_i) RVs with $p_i \ge 1/4$.*