Exercise 1 - Answers to Q4

1. Show that $\mathbb{E}|X| < \infty$

The assumption implies that for any $\epsilon \in (0, c)$, both $\mathbb{E}e^{\epsilon X}$ and $\mathbb{E}e^{-\epsilon X}$ are finite. As a result,

$$\mathbb{E}e^{\epsilon|X|} = \mathbb{E}e^{\epsilon X}1_{X\geq 0} + \mathbb{E}e^{-\epsilon X}1_{X<0} \leq \mathbb{E}e^{\epsilon X} + \mathbb{E}e^{-\epsilon X} < +\infty. \tag{1}$$

Choose a sufficiently large $M = M(\epsilon) > 0$ so that for all $|x| > M$, we have $e^{\epsilon|x|} \geq |x|$. It follows

$$\mathbb{E}|X| = \mathbb{E}[|X|1_{|X|\leq M}] + \mathbb{E}[|X|1_{|X|>M}]$$

By definition the first term is smaller than M. By our choice of M, the second term is smaller than $\mathbb{E}e^{\epsilon|X|}$, which by Eq. (1) is finite. In summary $\mathbb{E}|X| < \infty$.

2. Find the k-th derivative of $\psi(\lambda) = \mathbb{E}e^{\lambda X}$

The first derivative is by definition $\frac{d}{d\lambda}\mathbb{E}e^{\lambda X}$. If we could exchange order of differentiation w.r.t. λ and expectation w.r.t. X (e.g. integration w.r.t. its density), then the result would be $\mathbb{E}[Xe^{\lambda X}]$. Let us now formally prove that this indeed can be done. A similar result holds for higher order derivatives. Finally, note that at $\lambda = 0$, $\psi'(0) = \mathbb{E}X$ and similarly the k-th derivative $\psi^{(k)}(0) = \mathbb{E}[X^k]$, hence the name *moment generating function.*

To this end, recall the formal definition of the derivative

$$\frac{d}{d\lambda}\mathbb{E}e^{\lambda X} = \lim_{h\to 0} \frac{\mathbb{E}\left(e^{(\lambda+h)X} - e^{\lambda X}\right)}{h} = \lim_{h\to 0} \mathbb{E}\left(e^{\lambda X}\frac{e^{hX} - 1}{h}\right).$$

We will show that the limit exists by invoking the dominated convergence theorem.

Theorem 1. *Let $(Z_n)_{n\geq 1}$ be a sequence of real-valued random variables. Suppose that the sequence converges point-wise to another random variable Z_∞ and that there exists an integrable random variable Y such that $|Z_n| \leq Y$ for all n. Then Z_∞ is integrable and satisfies*

$$\lim_{n\to\infty} \mathbb{E}Z_n = \mathbb{E}Z_\infty.$$

Let us take an arbitrary sequence of real numbers $(h_n)_{n\geq 1}$ tending to zero. If

$$\sup_n \left| e^{\lambda X}\frac{e^{h_n X} - 1}{h_n} \right|$$

can be bounded by an integrable random variable, we can apply the theorem to show that the limit exists and is

$$\frac{d}{d\lambda}\mathbb{E}e^{\lambda X} = \mathbb{E}\left(Xe^{\lambda X}\right).$$

For a fixed $\lambda \in (-c, c)$, let us choose an $H > 0$ satisfying $|\lambda| + 2H < c$. Now for any $x \in \mathbb{R}$ and any $h \in (-H, H)$, we have

$$\left|\frac{e^{hx} - 1}{h}\right| \leq \left|\frac{1}{h}\int_0^h x e^{sx} ds\right| \leq |x|e^{|hx|} \leq |x|e^{H|x|}.$$

Again, we can choose an interval $[-M, M]$ outside of which the bound $|x|e^{H|x|} \leq e^{2H|x|}$ holds. It results in the following uniform bound

$$|x|e^{H|x|} \leq Me^{HM}1_{|x|\leq M} + e^{2H|x|}1_{|x|>M} \leq Me^{HM} + e^{2H|x|}.$$

Putting pieces together, we obtain

$$\sup_{|h|<H} e^{\lambda X}\left|\frac{e^{hX} - 1}{h}\right| \leq e^{|\lambda X|}(Me^{HM} + e^{2H|X|}).$$

This is the desired upper bound because in view of the previous question, we know $\mathbb{E}e^{(2H+|\lambda|)|X|} < +\infty$ and $\mathbb{E}e^{|\lambda X|} < +\infty$. Higher order derivatives can be proved in a similar manner.

3. Prove that $\mathbb{E}[e^{\lambda(X-\mathbb{E}X)}] \leq e^{s\lambda^2}$

Consider the function $g(\lambda) = \ln \mathbb{E}[e^{\lambda(X-\mathbb{E}X)}] = \ln \mathbb{E}[e^{\lambda X}] - \lambda \mathbb{E}X$. Following the previous question, let us take the first derivative

$$g'(\lambda) = \frac{\mathbb{E}[Xe^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]} - \mathbb{E}X.$$

As a result, we find that both $g(0) = g'(0) = 0$. Since its second derivative $g^{(2)}(\lambda)$ is continuous and finite in the vicinity of the origin, say, $[-c/2, c/2]$, we deduce, from a Taylor expansion of g around $\lambda = 0$ that there exists some constant s such that

$$g(\lambda) \leq s\lambda^2$$

for any $\lambda \in [-c/2, c/2]$. Hence the result.

4. Find an upper bound on $\mathbb{P}(X \geq \mathbb{E}X + t)$

For any $\lambda \in (0, c/2]$

$$\mathbb{P}(X \geq \mathbb{E}X + t) = \mathbb{P}(\lambda(X - \mathbb{E}X) \geq \lambda t) = \mathbb{P}(e^{\lambda(X-\mathbb{E}X)} \geq e^{\lambda t}) \leq \frac{\mathbb{E}[e^{\lambda(X-\mathbb{E}X)}]}{e^{\lambda t}} \leq e^{s\lambda^2 - \lambda t}$$

The first inequality is Markov's inequality, and the second is using the result of question 3. Since we want a tight bound, we'll minimize the exponent under constraint $\lambda \in \mathbb{R}_+$. Note that the exponent is a quadratic function of λ which achieves its minimum at

$$\lambda = \frac{t}{2s}.$$

If $t \in [0, sc]$, the minimum is attained in the interval $[0, c/2]$, which results in

$$\mathbb{P}(X \geq \mathbb{E}X + t) \leq e^{s\lambda^2 - \lambda t} = e^{-\frac{t^2}{4s}}$$

And if $t > sc$, the constraint is active, meaning that the minimum is achieved at $\lambda = \frac{c}{2}$

$$\mathbb{P}(X \geq \mathbb{E}X + t) \leq e^{s\lambda^2 - \lambda t} = e^{\frac{sc^2}{4} - \frac{ct}{2}} \leq e^{-\frac{ct}{4}}$$

where the last inequality uses the condition $t > sc$.

5. Prove a one-sided concentration inequality for χ_m^2

The expectation of a χ_m^2 distributed random variable Z is

$$\mathbb{E}Z = \mathbb{E}\left[\sum_{i=1}^{m} X_i^2\right] = \sum_{i=1}^{m} \mathbb{E}[X_i^2] = m.$$

We use $\mathbb{E}[X_i^4] = 3$ to calculate its second moment

$$\mathbb{E}Z^2 = \mathbb{E}\left[\sum_{i=1}^{m}\sum_{j=1}^{m} X_i^2 X_j^2\right] = \sum_{i=1}^{m} \mathbb{E}[X_i^4] + \sum_{i \neq j} \mathbb{E}[X_i^2 X_j^2] = 3m + m(m-1) = 2m + m^2.$$

Therefore, we deduce its variance

$$\mathbb{E}[Z^2] - \big(\mathbb{E}[Z]\big)^2 = 2m + m^2 - m^2 = 2m.$$

Now, to prove the desired concentration inequality, we calculate Z's moment generating function

$$\forall s \in (-\infty, 1/2), \quad \mathbb{E}e^{sZ} = (1-2s)^{-m/2}$$

in order to apply Markov's inequality

$$\mathbb{P}\big(Z^2 - m \geq 2m(t + \sqrt{t})\big) \leq \frac{\mathbb{E}e^{s(Z^2 - m)}}{e^{2sm(\sqrt{t}+t)}} = \frac{e^{-ms}(1-2s)^{-m/2}}{e^{2sm(\sqrt{t}+t)}}.$$

We can then optimize over $s \in (-\infty, 1/2)$ to obtain the desired bound

$$\mathbb{P}\big(Z^2 - m \geq 2m(t + \sqrt{t})\big) \leq \min_{s \in (-\infty, 1/2)} e^{-2sm(\sqrt{t}+t) - ms}(1-2s)^{-m/2} = \frac{(1 + 2t + 2\sqrt{t})^{m/2}}{e^{m(t+\sqrt{t})}} \leq e^{-mt}$$

where the minimum in the first inequality is achieved at

$$s^* = \frac{\sqrt{t} + t}{1 + 2(\sqrt{t} + t)} \in [0, 1/2).$$