EXERCISE 2

1. Prove Jensen's inequality for real-valued random variables: Let X be a real-valued random variable and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex. If $\mathbb{E}(X)$ and $\mathbb{E}(\varphi(X))$ are defined, then $\mathbb{E}(\varphi(X)) \geq \varphi(\mathbb{E}(X))$ with the convention $\varphi(+\infty) = \lim_{x\to\infty} \varphi(x)$ and $\varphi(-\infty) = \lim_{x\to-\infty} \varphi(x)$.
2. For $X_1 \sim \text{Poisson}(\theta)$, $\Lambda^*(x) = \theta - x + x\log\left(\frac{x}{\theta}\right)$ for nonnegative x, and $\Lambda^*(x) = \infty$ otherwise.
3. For $X_1 \sim \text{Bernoulli}(p)$, $\Lambda^*(x) = x\log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{1-x}{1-p}\right)$ for $x \in [0,1]$ and $\Lambda^*(x) = \infty$ otherwise. Note that $D_\Lambda = \mathbb{R}$, but Λ^* is discontinuous.
4. For $X_1 \sim \text{Exponential}(\theta)$, $\Lambda^*(x) = \theta x - 1 - \log(\theta x)$ for $x > 0$ and $\Lambda^*(x) = \infty$ otherwise.
5. For $X_1 \sim \text{Normal}(0, \sigma^2)$, $\Lambda^*(x) = \frac{x^2}{2\sigma^2}$.

Throughout the exercise, we let $g(x,\lambda) = \lambda x - \Lambda(\lambda)$ so that $\partial_\lambda g = x - \Lambda'(\lambda)$.

1. We make use the property of convex functions that

$$\varphi(x) = \sup\{ax + b : \varphi(x) \geq ax + b \text{ for all } x \in \mathbb{R}\}. \tag{0.1}$$

The left-hand side is by definition larger than the other, and thus it remains to show the remaining. To this end, it suffices to show that for all (x_0, y_0) satisfying $\varphi(x_0) \geq y_0$, there exists a linear function $\psi(x) = a(x - x_0) + y_0$ such that $\varphi(x) \leq \psi(x)$ for all x. Essentially, this is achieved by choosing

$$\begin{aligned} a &\in \left[\sup_{x<x_0} \frac{\varphi(x) - \varphi(x_0)}{x - x_0}, \inf_{x>x_0} \frac{\varphi(x) - \varphi(x_0)}{x - x_0}\right] \\ &= \left[\liminf_{x\to x_0} \frac{\varphi(x) - \varphi(x_0)}{x - x_0}, \limsup_{x\to x_0} \frac{\varphi(x) - \varphi(x_0)}{x - x_0}\right]. \end{aligned}$$

Now that (0.1) coincides with

$$\varphi(x) = \sup\{ax + b : \varphi(x) \geq ax + b \text{ for all } x \in \mathbb{R}, a, b \in \mathbb{Q}\}, \tag{0.2}$$

one can enumerate the linear functions on the right-hand side by $(\psi_i)_{i=1}^\infty$ and obtain

$$\mathbb{E}(\varphi(X)) \geq \mathbb{E}\left(\max_{1\leq i\leq n} \psi_i(X)\right) \geq \max_{1\leq i\leq n} \psi_i(\mathbb{E}(X)).$$

If $\mathbb{E}(X) \in \mathbb{R}$, then the proposition holds by letting $n \to \infty$. If $\mathbb{E}(X) = \infty$ (similar for $\mathbb{E}(X) = -\infty$) and $\varphi(\infty) > -\infty$, then right-hand side of the above still converges to $\varphi(\infty)$, while the proposition is trivial when $\mathbb{E}(X) = \infty$ and $\varphi(\infty) = -\infty$.

2. Note that $\mathbb{P}(X_1 = k) = \frac{\theta^k e^{-\theta}}{k!}$ and that

$$\Lambda(\lambda) = \log \sum_{i=0}^\infty \frac{\theta^k e^{-\theta}}{k!} \cdot e^{\lambda k} = \log \sum_{i=0}^\infty \frac{\left(\theta e^\lambda\right)^k e^{-\theta}}{k!} = \theta e^\lambda - \theta.$$

To derive $\Lambda^*(x)$, we calculate the derivative:

$$\partial_\lambda g = x - \Lambda'(\lambda) = x - \theta e^\lambda.$$

By monotonicity of $\Lambda'(\lambda)$, we derive $\Lambda^*(x)$ in the following cases.

(a) If $x \leq 0$, then $\partial_\lambda g \leq 0$ and hence

$$\Lambda^*(x) = \lim_{\lambda\to-\infty} g(x,\lambda) = \begin{cases} -\theta & \text{if } x = 0, \\ \infty & \text{if } x < 0. \end{cases}$$

(b) If $x > 0$, then there exists $\lambda \in \mathbb{R}$ such that $x = \Lambda'(\lambda)$ (i.e., $\lambda = \log(\frac{x}{\theta})$), and hence

$$\Lambda^*(x) = g(x,\lambda) = \theta - x + x\log\left(\frac{x}{\theta}\right).$$

3. Straightforwardly, $\Lambda(\lambda) = \log\bigl(pe^\lambda + (1-p)\bigr)$ and

$$\partial_\lambda g = x - \Lambda'(\lambda) = x - \frac{pe^\lambda}{pe^\lambda + (1-p)}.$$

By monotonicity of $\Lambda'(\lambda)$, we have the following.

(a) If $x \leq 0$, then $\partial_\lambda g \leq 0$ and hence

$$\Lambda^*(x) = \lim_{\lambda \to -\infty} g(x, \lambda) = \begin{cases} -\log(1-p) & \text{if } x = 0, \\ \infty & \text{if } x < 0. \end{cases}$$

(b) If $x \geq 1$, then $\partial_\lambda g \geq 0$ and hence

$$\Lambda^*(x) = \lim_{\lambda \to \infty} g(x, \lambda) = \begin{cases} -\log p & \text{if } x = 1, \\ \infty & \text{if } x < 0. \end{cases}$$

(c) If $x \in (0,1)$, then there exists $\lambda \in \mathbb{R}$ such that $x = \Lambda'(\lambda)$ (i.e., $\lambda = \log \frac{(1-p)x}{p(1-x)}$ and $\Lambda(\lambda) = \log \frac{1-p}{1-x}$), and hence

$$\Lambda^*(x) = g(x, \lambda) = x \log\left(\frac{x}{p}\right) + (1-x) \log\left(\frac{1-x}{1-p}\right).$$

4. Note that $X_1 \sim \theta e^{-\theta x} \mathbb{1}_{[0,\infty)} dx$ and that

$$\Lambda(\lambda) = \log \int_0^\infty e^{\lambda x} \theta e^{-\theta x} dx = \begin{cases} \infty & \text{if } \ \lambda \geq \theta, \\ \log \theta - \log(\theta - \lambda) \end{cases} \text{if } \ \lambda < \theta.)$$

and that

$$\partial_\lambda g = x - \Lambda'(\lambda) = x - \frac{1}{\theta - \lambda}.$$

By monotonicity of $\Lambda'(\lambda)$, we have the following.
(a) If $x \leq 0$, then $\partial_\lambda g \leq 0$ for $\lambda < \theta$ and hence

$$\Lambda^*(x) = \lim_{\lambda \to -\infty} g(x, \lambda) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x < 0. \end{cases}$$

(b) If $x \in (0, \infty)$, then there exists $\lambda < \theta$ such that $x = \Lambda'(\lambda)$ (i.e., $\lambda = \theta - \frac{1}{x}$), and hence

$$\Lambda^*(x) = g(x, \lambda) = \theta x - 1 - \log(\theta x).$$

5. Note that $X_1 \sim \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx$ and that

$$\Lambda(\lambda) = \log \int_{-\infty}^\infty e^{\lambda x} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{\sigma^2 \lambda^2}{2} + \log \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(x - \sigma^2 \lambda\right)^2}{2\sigma^2}} dx = \frac{\sigma^2 \lambda^2}{2}$$

and that

$$\partial_\lambda g = x - \Lambda'(\lambda) = x - \sigma^2 \lambda.$$

Hence, for all $x \in \mathbb{R}$ there exists $\lambda \in \mathbb{R}$ such that $x = \Lambda'(\lambda)$ (i.e., $\lambda = \frac{x}{\sigma^2}$) and that

$$\Lambda^*(x) = g(x, \lambda) = \frac{x^2}{2\sigma^2}.$$