Limit/large dev. thms. HW assignment 1.

1. (a) Express the logarithmic moment generating function (see page 7 of scanned lecture notes) of $aX+b$ in terms of the logarithmic moment generating function of X.

   (b) Let X and Y denote independent random variables. Express the logarithmic moment generating function of $X+Y$ in terms of the logarithmic moment generating functions of X and Y.

   (c) Let $X_1, X_2, \dots$ denote i.i.d. random variables and let N denote a non-negative integer-valued random variable, which is independent from $X_1, X_2, \dots$ Let

$$Y = X_1 + \dots X_N.$$

   Denote by $\widehat{I}$ the log. mom. gen. function of X_i and denote by $\widehat{J}$ the log. mom. gen. function of N. Show that the log. mom. gen. function of Y is $\widehat{J} \circ \widehat{I}$.

Solution:

(a) $Z_X(\lambda) = \mathbb{E}[e^{\lambda X}]$. Let $Y = aX+b$. We have

$$Z_Y(\lambda) = \mathbb{E}[e^{\lambda Y}] = \mathbb{E}[e^{\lambda a X + \lambda b}] = e^{\lambda b}\mathbb{E}[e^{\lambda a X}] = e^{\lambda b} Z(a\lambda),$$

thus $\widehat{I}_Y(\lambda) = \ln(Z_Y(\lambda)) = \lambda b + \widehat{I}_X(a\lambda)$.

(b) Let $Z = X+Y$. Then we have

$$\widehat{I}_Z(\lambda) = \ln(\mathbb{E}[e^{\lambda Z}]) = \ln(\mathbb{E}[e^{\lambda X}e^{\lambda Y}]) \overset{(*)}{=} \ln(\mathbb{E}[e^{\lambda X}]\mathbb{E}[e^{\lambda Y}]) = \widehat{I}_X(\lambda) + \widehat{I}_Y(\lambda),$$

where $(*)$ follows from the fact that the expectation of the product of independent random variables is equal to the product of their expectations.

(c) Let $p_k = \mathbb{P}(N=k)$ for $k = 0, 1, 2, \dots$
Then $\widehat{J}(\lambda) = \ln\left(\mathbb{E}[e^{\lambda N}]\right) = \ln\left(\sum_{k=0}^{\infty} p_k e^{\lambda k}\right)$.
Let $Z(\lambda) = \mathbb{E}(e^{\lambda X_i})$, thus $\widehat{I}(\lambda) = \ln(Z(\lambda))$.

$$Z_Y(\lambda) = \mathbb{E}[e^{\lambda Y}] = \mathbb{E}[e^{\lambda X_1 + \dots + \lambda X_N}] = \sum_{k=0}^{\infty} \mathbb{E}[e^{\lambda X_1 + \dots + \lambda X_N} \mid N = k]\mathbb{P}(N=k) =$$

$$\sum_{k=0}^{\infty} \mathbb{E}[e^{\lambda X_1} \dots e^{\lambda X_k}] \cdot p_k = \sum_{k=0}^{\infty} \mathbb{E}[e^{\lambda X_1}] \dots \mathbb{E}[e^{\lambda X_k}] \cdot p_k = \sum_{k=0}^{\infty} p_k Z(\lambda)^k,$$

thus

$$\widetilde{I}_Y(\lambda) = \ln(Z_Y(\lambda)) = \ln\left(\sum_{k=0}^{\infty} p_k Z(\lambda)^k\right) = \ln\left(\sum_{k=0}^{\infty} p_k e^{\widehat{I}(\lambda)k}\right) = \widehat{J}\left(\widehat{I}(\lambda)\right).$$

2. Let $Y \sim \mathrm{POI}(10000)$ (Poisson distribution with parameter 10000). The goal of this exercise is to estimate the number of zero digits (after the decimal point) before the first non-zero digit in the decimal expansion of the probability $\mathbb{P}(Y \geq 27182)$. You will give an upper bound and a lower bound using different methods.

   (a) Calculate the logarithmic moment generating function $\widehat{I}(\lambda)$ of the $\mathrm{POI}(\mu)$ distribution (see page 7 of the scanned lecture notes) and calculate its Legendre transform $I(x)$ (page 9 of scanned).

   (b) In order to give an upper bound on $\mathbb{P}(Y \geq 27182)$, use the *exponential Chebyshev's inequality* (i.e., the method that we used on the top of page 8 of the scanned lecture notes).

   (c) In order to give a lower bound on $\mathbb{P}(Y \geq 27182)$, estimate $\mathbb{P}(Y = 27182)$ using the crude version of Stirling's formula (page 3 of scanned).

   (d) Based on the above calculations, what is the approximate number of zero digits (after the decimal point) before the first non-zero digit in the decimal expansion of the probability $\mathbb{P}(Y \geq 27182)$?

Solution:

(a) If $X \sim \mathrm{POI}(\mu)$, then $\mathbb{P}(X = k) = e^{-\mu}\frac{\mu^k}{k!}$, hence

$$Z(\lambda) = \mathbb{E}(e^{\lambda X}) = \sum_{k=0}^{\infty} e^{\lambda k} e^{-\mu} \frac{\mu^k}{k!} = e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu e^{\lambda})^k}{k!} = e^{-\mu} \exp(\mu e^{\lambda}) = \exp(\mu \cdot (e^{\lambda} - 1)),$$

which implies that $\widehat{I}(\lambda) = \ln(Z(\lambda)) = \mu \cdot (e^{\lambda} - 1)$. Now $I(x) = \max_{\lambda \in \mathbb{R}}\{x\lambda - \widehat{I}(\lambda)\}$, thus we first need to find $\lambda^* = \lambda^*(x)$ such that $x = \widehat{I}'(\lambda^*)$. Now $\widehat{I}'(\lambda) = \mu e^{\lambda}$, thus $\lambda^* = \ln(\frac{x}{\mu})$ and

$$I(x) = x\lambda^* - \widehat{I}(\lambda^*) = x \ln(\frac{x}{\mu}) - \mu \cdot (e^{\ln(\frac{x}{\mu})} - 1) = x \ln\left(\frac{x}{\mu}\right) + \mu - x \quad \text{if} \quad x \geq 0.$$

Note that if $x < 0$ then $I(x) = +\infty$ because $\lim_{\lambda \to -\infty}(x\lambda - \widehat{I}(\lambda)) = +\infty$.

(b) $27182 \approx 10000 \cdot e$, thus $\mu = 10000$, $x = 10000 \cdot e$ and $\lambda^* = \ln(\frac{x}{\mu}) = \ln(e) = 1$ and

$$\mathbb{P}(Y \geq 27182) = \mathbb{P}(e^{\lambda^* Y} \geq e^{\lambda^* 27182}) \leq \frac{\mathbb{E}(e^{\lambda^* Y})}{e^{\lambda^* 27182}} = \frac{\mathbb{E}(e^{Y})}{e^{10000e}} = \frac{\exp(10000 \cdot (e-1))}{e^{10000e}} = e^{-10000}$$

(c) $\mathbb{P}(Y \geq 27182) \geq \mathbb{P}(Y = 27182) = e^{-10000}\frac{10000^{27182}}{27182!}$.

Now we crudely replace $27182!$ by $27182^{27182}e^{-27182}$, so $e^{-10000}\frac{10000^{27182}}{27182!}$ is crudely replaced by

$$e^{-10000}\frac{10000^{27182}}{27182^{27182}e^{-27182}} \stackrel{(*)}{=} e^{-10000}\frac{10000^{27182}}{10000^{27182} \cdot e^{27182}e^{-27182}} = e^{-10000},$$

where in $(*)$ we also replaced 21782 by $10000 \cdot e$. Of course this calculation was not entirely rigorous: in order to make it rigorous, we can use more precise versions of Stirling's formula.

(d) We see from the upper bound of (b) and (non-rigorous) lower bound of (c) that it is OK to replace $\log_{10}(\mathbb{P}(Y \geq 2718))$ by $\log_{10}(e^{-10000}) = -10000 \cdot \log_{10}(e) \approx -4343$.

Thus the number of zero digits before the first non-zero digit in the decimal expansion of the probability $\mathbb{P}(Y \geq 27182)$ is roughly 4343.

Remark: This exercise can be viewed as a large deviation theorem for the sum of i.i.d. random variables. If $X_1, X_2, \dots$ are i.i.d. with $\mathrm{POI}(\mu)$ distribution and $S_n = X_1 + \dots + X_n$, then $S_n \sim \mathrm{POI}(n\mu)$. So what we have just proved is a special case of *Cramér's theorem*, which implies that for any $x \geq \mu$ we have

$$\lim_{n \to \infty} \frac{1}{n} \ln\left(\frac{S_n}{n} \geq x\right) = -I(x),$$

where $I(x)$ was calculated in part (a) of the exercise. What we estimated in parts (b) and (c) amounts to the case $\mu = 1$, $n = 10000$ and $x = e$.

3. *Laplace's principle.* Let $-\infty \le a < b \le +\infty$ and let $J : (a,b) \to \mathbb{R}$ denote a continuous function. Let us also assume that there is $x^* \in (a,b)$ for which $J(x^*) = \min_{x\in(a,b)} J(x)$ and that $\int_a^b e^{-J(x)}\,\mathrm{d}x < +\infty$. Prove that

$$\lim_{n\to\infty} -\frac{1}{n}\ln\left(\int_a^b e^{-nJ(x)}\,\mathrm{d}x\right) = J(x^*). \tag{1}$$

Hint: Prove the liminf bound and the limsup bound separately.

Solution: Let us denote

$$\alpha = J(x^*) = \min_{x\in(a,b)} J(x).$$

We have

$$e^{-nJ(x)} \le e^{-(n-1)\alpha}e^{-J(x)}, \qquad x \in (a,b),$$

thus

$$\ln\left(\int_a^b e^{-nJ(x)}\,\mathrm{d}x\right) \le \ln\left(e^{-(n-1)\alpha}\int_a^b e^{-J(x)}\,\mathrm{d}x\right) = -(n-1)\alpha + \ln\left(\int_a^b e^{-J(x)}\,\mathrm{d}x\right),$$

thus

$$\limsup_{n\to\infty}\frac{1}{n}\ln\left(\int_a^b e^{-nJ(x)}\,\mathrm{d}x\right) \le \lim_{n\to\infty}\frac{1}{n}\left(-(n-1)\alpha + \ln\left(\int_a^b e^{-J(x)}\,\mathrm{d}x\right)\right) = -\alpha. \tag{2}$$

Now we want to bound the integral in the other direction. We will show that for any $\varepsilon > 0$ we have

$$\liminf_{n\to\infty}\frac{1}{n}\ln\left(\int_a^b e^{-nJ(x)}\,\mathrm{d}x\right) \ge -(\alpha+\varepsilon). \tag{3}$$

Note that the fact that (3) holds for any $\varepsilon > 0$, together with (2), implies (1). It remains to show (3).

Let us fix $\varepsilon > 0$. Taking into account that J is continuous, we can find $\delta > 0$ such that for any $x \in [x^*-\delta, x^*+\delta]$ we have $J(x) \le \alpha + \varepsilon$. Therefore we have

$$\int_a^b e^{-nJ(x)}\,\mathrm{d}x \ge \int_{x^*-\delta}^{x^*+\delta} e^{-nJ(x)}\,\mathrm{d}x \ge \int_{x^*-\delta}^{x^*+\delta} e^{-n(\alpha+\varepsilon)}\,\mathrm{d}x = 2\delta e^{-n(\alpha+\varepsilon)},$$

therefore

$$\liminf_{n\to\infty}\frac{1}{n}\ln\left(\int_a^b e^{-nJ(x)}\,\mathrm{d}x\right) \ge \lim_{n\to\infty}\frac{1}{n}\ln\left(2\delta e^{-n(\alpha+\varepsilon)}\right) = -(\alpha+\varepsilon).$$

This proves (3).