Limit/large dev. thms. HW assignment 2., Solutions

1. Let $\lambda \mapsto Z(\lambda)$ denote the moment generating function of the random variable X. If $Z(\mu) < +\infty$, let $X^{(\mu)}$ denote the *exponentially tilted* random variable. The cumulative distribution function of $X^{(\mu)}$ is

$$F_\mu(x) = \mathbb{P}(X^{(\mu)} \le x) = \frac{\mathbb{E}(e^{\mu X}\mathbb{1}[X \le x])}{Z(\mu)}.$$

We have learnt some facts about exponential tilting in class, see page 17-18 of scanned lecture notes.

(a) Show that ${X^{(\lambda)}}^{(\mu)} \sim X^{(\lambda+\mu)}$. In words: tilting the tilted random variable amounts to tilting the original random variable with the sum of the two tiltings.

(b) Express the logarithmic moment generating function $\widehat{I}_\mu$ of $X^{(\mu)}$ using the logarithmic moment generating function $\widehat{I}$ of X.

(c) Express the Legendre transform I_μ of $\widehat{I}_\mu$ using the Legendre transform I of $\widehat{I}$.

Solution:

(a) We have learnt in class (see page 18 of scanned) that $\mathbb{E}(g(X^{(\mu)})) = \mathbb{E}(e^{\mu X}g(X))/\mathbb{E}(e^{\mu X})$. We apply this formula in the numerator as well as the denominator in equation $(*)$ below:

$$\mathbb{P}\left({X^{(\mu)}}^{(\lambda)} \le x\right) = \frac{\mathbb{E}\left(e^{\lambda X^{(\mu)}}\mathbb{1}[X^{(\mu)} \le x]\right)}{\mathbb{E}(e^{\lambda X^{(\mu)}})} \overset{(*)}{=} \frac{\mathbb{E}\left(e^{\mu X}e^{\lambda X}\mathbb{1}[X \le x]\right)/\mathbb{E}\left(e^{\mu X}\right)}{\mathbb{E}(e^{\mu X}e^{\lambda X})/\mathbb{E}\left(e^{\mu X}\right)} =$$
$$\frac{\mathbb{E}\left(e^{(\lambda+\mu)X}\mathbb{1}[X \le x]\right)}{\mathbb{E}(e^{(\lambda+\mu)X})} = \mathbb{P}[X^{(\lambda+\mu)} \le x]$$

This shows that ${X^{(\mu)}}^{(\lambda)} \sim X^{(\mu+\lambda)}$.

(b) $\widehat{I}_\mu(\lambda) = \ln\left(\mathbb{E}(e^{\lambda X^{(\mu)}})\right) = \ln\left(\mathbb{E}(e^{\mu X}e^{\lambda X})/\mathbb{E}(e^{\mu X})\right) = \widehat{I}(\mu+\lambda) - \widehat{I}(\mu)$

(c)

$$I_\mu(x) = \sup_\lambda\{\lambda x - \widehat{I}_\mu(\lambda)\} = \sup_\lambda\{\lambda x - \widehat{I}(\mu+\lambda) + \widehat{I}(\mu)\} =$$
$$\widehat{I}(\mu) - \mu x + \sup_\lambda\{(\lambda+\mu)x - \widehat{I}(\mu+\lambda)\} = \widehat{I}(\mu) - \mu x + I(x) = I(x) - \mu x + \widehat{I}(\mu).$$

Remark: Alternative proof of $I_\mu(x) = I(x) - \mu x + \widehat{I}(\mu)$ using our non-rigorous Cramér formalism: we know that $X_1^{(\mu)} + \cdots + X_n^{(\mu)}$ has the same distribution as $S_n^{(\mu)}$, thus

$$e^{-nI_\mu(x)} \approxeq \mathbb{P}\left(S_n^{(\mu)} \approx nx\right) = \frac{e^{\mu nx}}{\mathbb{E}(e^{\lambda S_n})}\mathbb{P}\left(S_n \approx nx\right) \approxeq \frac{e^{\mu nx}}{Z(\lambda)^n}e^{-nI(x)} = e^{-n(I(x)-\mu x+\widehat{I}(\mu))}.$$

2. In this exercise $X^{(\mu)}$ denotes the random variable that we obtain by exponentially tilting the distribution of the random variable X.

   (a) Show that if $X \sim \text{BIN}(n,p)$ then $X^{(\mu)} \sim \text{BIN}(n,p')$ for some $p' = p'(p,\mu)$ and that any $p' \in (0,1)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

   (b) Show that if $X \sim \text{POI}(\lambda)$ then $X^{(\mu)} \sim \text{POI}(\lambda')$ for some $\lambda' = \lambda'(\lambda,\mu)$ and that any $\lambda' \in (0,+\infty)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

   (c) If $X \sim \text{EXP}(\lambda)$, find the values of μ for which $Z(\mu) < +\infty$ (i.e., find the domain of the moment generating function $Z(\cdot)$). Show that $X^{(\mu)} \sim \text{EXP}(\lambda')$ for some $\lambda' = \lambda'(\lambda,\mu)$ and that any $\lambda' \in (0,+\infty)$ can be obtained by choosing μ from the domain of $Z(\cdot)$ appropriately.

   (d) Show that if $X \sim \mathcal{N}(m,\sigma^2)$ then $X^{(\mu)} \sim \mathcal{N}(m',\sigma^2)$ for some $m' = m'(m,\mu,\sigma)$ and that any $m' \in (-\infty,+\infty)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

Solution:

(a) If $X \sim \text{BIN}(n,p)$ then $Z(\mu) = (pe^\mu + (1-p))^n$ (since X is the sum of n i.i.d. $\text{BER}(p)$ random variables), thus

$$\mathbb{P}(X^{(\mu)} = k) = (pe^\mu + (1-p))^{-n} e^{\mu k} \binom{n}{k} p^k (1-p)^{n-k} =$$

$$\binom{n}{k} \left(\frac{pe^\mu}{pe^\mu + (1-p)} \right)^k \left(\frac{(1-p)}{pe^\mu + (1-p)} \right)^{n-k}, \quad \text{thus} \quad X^{(\mu)} \sim \text{BIN}\left(n, \frac{pe^\mu}{pe^\mu + (1-p)}\right).$$

Thus $p'(p,\mu) = \frac{pe^\mu}{pe^\mu + (1-p)}$. Clearly, $\lim_{\mu \to -\infty} p'(p,\mu) = 0$ and $\lim_{\mu \to +\infty} p'(p,\mu) = 1$, so indeed any $p' \in (0,1)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

(b) If $X \sim \text{POI}(\lambda)$ then $Z(\mu) = e^{(e^\mu - 1)\lambda}$, thus

$$\mathbb{P}(X^{(\mu)} = k) = e^{(1-e^\mu)\lambda} e^{\mu k} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-e^\mu \lambda} \frac{(e^\mu \lambda)^k}{k!}, \quad \text{thus} \quad X^{(\mu)} \sim \text{POI}(e^\mu \lambda).$$

Thus $\lambda'(\lambda,\mu) = e^\mu \lambda$. Clearly, $\lim_{\mu \to -\infty} \lambda'(\lambda,\mu) = 0$ and $\lim_{\mu \to +\infty} \lambda'(\lambda,\mu) = +\infty$, so indeed any $\lambda' \in (0,+\infty)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

(c) If $X \sim \text{EXP}(\lambda)$ then the density function of X is $f(x) = \lambda e^{-\lambda x} \mathbb{1}[x \geq 0]$, thus

$$Z(\mu) = \mathbb{E}(e^{\mu X}) = \int_0^\infty e^{\mu x} f(x)\,\mathrm{d}x = \int_0^\infty \lambda e^{(\mu - \lambda)x}\,\mathrm{d}x = \frac{\lambda}{\lambda - \mu}, \quad \text{if} \quad \mu < \lambda.$$

If $\mu \geq \lambda$ then $Z(\mu) = +\infty$.
Thus $f^{(\mu)}(x) = Z(\mu)^{-1} e^{\mu x} f(x) = (\lambda - \mu) e^{-(\lambda - \mu)x} \mathbb{1}[x \geq 0]$, thus $X^{(\mu)} \sim \text{EXP}(\lambda - \mu)$, if $\mu < \lambda$. Thus $\lambda'(\lambda,\mu) = \lambda - \mu$. Clearly, any $\lambda' \in (0,+\infty)$ can be obtained by choosing $\mu = \lambda - \lambda'$.

(d) If $X \sim \mathcal{N}(m,\sigma^2)$ then $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$ and $Z(\mu) = \exp(\mu m + \frac{1}{2}\sigma^2\mu^2)$, thus

$$f^{(\mu)}(x) = Z(\mu)^{-1} e^{\mu x} f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2} + \mu x - \mu m - \frac{1}{2}\sigma^2\mu^2\right) =$$

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - m - \sigma^2\mu)^2}{2\sigma^2}\right), \quad \text{thus} \quad X^{(\mu)} \sim \mathcal{N}\left(m + \sigma^2\mu, \sigma^2\right).$$

Thus, $m'(m,\mu,\sigma) = m + \sigma^2\mu$ and indeed any $m' \in (-\infty,+\infty)$ can be obtained by choosing $\mu \in \mathbb{R}$ appropriately.

3. Let $X_1, X_2, \ldots$ denote i.i.d. non-negative integer-valued random variables with distribution $\mathbb{P}(X_i = k) = p_k$, where $k = 0, 1, 2, \ldots$ Let $\lambda \in \mathbb{R}$ such that $Z(\lambda) = \mathbb{E}[e^{\lambda X_i}] < +\infty$. Let $S_n = X_1 + \cdots + X_n$.

Let $X_1^{(\lambda)}, X_2^{(\lambda)}, \ldots$ denote i.i.d. non-negative integer-valued random variables with distribution

$$\mathbb{P}(X_i^{(\lambda)} = k) = \frac{1}{Z(\lambda)} e^{\lambda k} p_k, \quad \text{where} \quad k = 0, 1, 2, \ldots$$

(a) Show that we have $\mathbb{P}(X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = k) = \frac{e^{\lambda k}\mathbb{P}(X_1 + \cdots + X_n = k)}{Z(\lambda)^n}$.

Instruction: This could be easily derived from the Lemma on page 20 of the scanned lecture notes, but since we only proved that lemma in the absolutely continuous case, I ask you to write down a complete proof of this sub-exercise only using the basic facts about exponential tilting (page 17-18 of scanned lecture notes).

(b) Show that $\mathbb{P}(X_1^{(\lambda)} = k \,|\, X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = m) = \mathbb{P}(X_1 = k \,|\, X_1 + \cdots + X_n = m)$.

(c) If $X_i \sim \mathrm{POI}(\mu)$, what is the conditional distribution of X_1 given that $X_1 + \cdots + X_n = \lfloor nx \rfloor$?

(d) If $X_i \sim \mathrm{POI}(\mu)$, show that for any $x > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}(X_1 = k \,|\, X_1 + \cdots + X_n = \lfloor nx \rfloor) = e^{-x}\frac{x^k}{k!}.$$

Remark: This last result is a rigorous version of the result proved heuristically on page 25 of the scanned lecture notes (also note that an exponentially tilted Poisson random variable is still a Poisson random variable with a different parameter)

Solution:

(a) Denote by Λ the set of n-tuples $(k_1, \ldots, k_n)$ of non-negative integers that satisfy $k_1 + \cdots + k_n = k$.

$$\mathbb{P}(X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = k) = \sum_{(k_1,\ldots,k_n)\in\Lambda} \mathbb{P}(X_1^{(\lambda)} = k_1, \ldots, X_n^{(\lambda)} = k_n) =$$

$$\sum_{(k_1,\ldots,k_n)\in\Lambda} \mathbb{P}(X_1^{(\lambda)} = k_1) \ldots \mathbb{P}(X_n^{(\lambda)} = k_n) = \sum_{(k_1,\ldots,k_n)\in\Lambda} \frac{e^{\lambda k_1}p_{k_1}}{Z(\lambda)} \cdots \frac{e^{\lambda k_n}p_{k_n}}{Z(\lambda)} =$$

$$\frac{e^{\lambda k}}{Z(\lambda)^n} \sum_{(k_1,\ldots,k_n)\in\Lambda} p_{k_1}\ldots p_{k_n} = \frac{e^{\lambda k}}{Z(\lambda)^n} \sum_{(k_1,\ldots,k_n)\in\Lambda} \mathbb{P}(X_1 = k_1, \ldots, X_n = k_n) =$$

$$\frac{e^{\lambda k}\mathbb{P}(X_1 + \cdots + X_n = k)}{Z(\lambda)^n}$$

(b)

$$\mathbb{P}(X_1^{(\lambda)} = k \,|\, X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = m) = \frac{\mathbb{P}(X_1^{(\lambda)} = k,\, X_2^{(\lambda)} + \cdots + X_n^{(\lambda)} = m - k)}{\mathbb{P}(X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = m)} =$$

$$\frac{\mathbb{P}(X_1^{(\lambda)} = k)\mathbb{P}(X_2^{(\lambda)} + \cdots + X_n^{(\lambda)} = m - k)}{\mathbb{P}(X_1^{(\lambda)} + \cdots + X_n^{(\lambda)} = m)} = \frac{\frac{e^{\lambda k}p_k}{Z(\lambda)} \frac{e^{\lambda(m-k)}\mathbb{P}(X_2+\cdots+X_n=m-k)}{(Z(\lambda))^{n-1}}}{\frac{e^{\lambda m}\mathbb{P}(X_1+\cdots+X_n=m)}{(Z(\lambda))^n}} =$$

$$\frac{p_k \cdot \mathbb{P}(X_2 + \cdots + X_n = m - k)}{\mathbb{P}(X_1 + \cdots + X_n = m)} = \mathbb{P}(X_1 = k \,|\, X_1 + \cdots + X_n = m)$$

(c) Let $Y = X_2 + \cdots + X_n$. Then $Y \sim \mathrm{POI}((n-1)\mu)$. For any $0 \le k \le \lfloor nx \rfloor$ we have

$$\mathbb{P}(X_1 = k \,|\, X_1 + Y = \lfloor nx \rfloor) = \frac{\mathbb{P}(X_1 = k,\, Y = \lfloor nx \rfloor - k)}{\mathbb{P}(X_1 + Y = \lfloor nx \rfloor)} = \frac{e^{-\mu}\frac{\mu^k}{k!} \cdot e^{-(n-1)\mu}\frac{((n-1)\mu)^{\lfloor nx \rfloor - k}}{(\lfloor nx \rfloor - k)!}}{e^{-n\mu}\frac{(n\mu)^{\lfloor nx \rfloor}}{\lfloor nx \rfloor!}} =$$

$$\frac{\frac{\mu^k}{k!} \cdot \frac{((n-1)\mu)^{\lfloor nx \rfloor - k}}{(\lfloor nx \rfloor - k)!}}{\frac{(n\mu)^{\lfloor nx \rfloor}}{\lfloor nx \rfloor!}} = \frac{\lfloor nx \rfloor!}{k!(\lfloor nx \rfloor - k)!} \frac{\mu^k \cdot ((n-1)\mu)^{\lfloor nx \rfloor - k}}{(n\mu)^{\lfloor nx \rfloor}} = \binom{\lfloor nx \rfloor}{k}\left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{\lfloor nx \rfloor - k},$$

thus the conditional distribution of X_1 given that $X_1 + Y = \lfloor nx \rfloor$ is $\mathrm{BIN}\left(\lfloor nx \rfloor, \frac{1}{n}\right)$.

(d)

$$\lim_{n\to\infty} \frac{\lfloor nx \rfloor !}{k!(\lfloor nx \rfloor - k)!} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{\lfloor nx \rfloor - k} =$$

$$\frac{1}{k!} \lim_{n\to\infty} \frac{\lfloor nx \rfloor}{n} \frac{\lfloor nx \rfloor - 1}{n} \cdots \frac{\lfloor nx \rfloor - (k-1)}{n} \left(1 - \frac{1}{n}\right)^{\lfloor nx \rfloor - k} = \frac{1}{k!} x^k e^{-x}.$$