Midterm Exam - May 3, 2024, Limit thms. of probab.

1. (8 points) Let X be a random variable with distribution $\mathbb{P}(X=k)=\frac{1}{e}\frac{1}{(k-1)!}$, $k=1,2,3,\dots$

   Let $X_1, X_2, \dots$ denote i.i.d. random variables with the same distribution as X. Let us define
   $$S_n = X_1 + \cdots + X_n.$$
   (a) Find the logarithmic moment generating function $\lambda \mapsto \ln(M(\lambda))$ of X.

   (b) Find the tilting parameter $\lambda_3 \in \mathbb{R}$ such that the exponentially tilted random variable $X^{(\lambda_3)}$ has expectation equal to 3.

   (c) Find the limit $R_3 = \lim_{n\to\infty}\frac{1}{n}\ln\left(\mathbb{P}\left[S_n \geq 3n\right]\right)$.

   (d) What is the relation between the values of $\ln(M(\lambda_3))$, λ_3 and R_3 according to Cramér's theorem? Check that this identity between the numbers that you found in (a),(b),(c) above indeed holds.

   Solution: Let $\widetilde{X} = X-1$, $\widetilde{X}_i := X_i - 1$, $\widetilde{S}_n = S_n - n = \widetilde{X_1}+\cdots+\widetilde{X}_n$. Note that $\widetilde{X} \sim \mathrm{POI}(1)$.

   (a) $\ln(M(\lambda)) = \ln\left(\mathbb{E}[e^{\lambda X}]\right) = \ln\left(e^{\lambda}\mathbb{E}[e^{\lambda\widetilde{X}}]\right) = \lambda + (e^{\lambda}-1)$, since $e^{\lambda}-1$ is the log.mom.gen. function of the POI(1) distribution by the formula sheet.

   (b) $X^{(\lambda)} = \widetilde{X}^{(\lambda)}+1$, so we want λ_3 such that $\mathbb{E}[\widetilde{X}^{(\lambda_3)}]=2$. We know from the formula sheet that $\widetilde{X}^{(\lambda)} \sim \mathrm{POI}(e^{\lambda})$, thus $\mathbb{E}[\widetilde{X}^{(\lambda)}]=e^{\lambda}$, thus $\lambda_3 = \ln(2)$.

   (c) $R_3 = \lim_{n\to\infty}\frac{1}{n}\ln\left(\mathbb{P}\left[\frac{\widetilde{S}_n}{n}\geq 2\right]\right) = -(2\ln(2)-1)$ by the formula sheet (the large dev. rate function of POI(1) distribution evaluated at 2).

   (d) We have $I(3) = 3\lambda^*(3) - \widehat{I}(\lambda^*(3))$, because I (the large dev. rate function of X) is the Legendre transform of $\widehat{I}$ (the log.mom.gen.fn. of X). With our notation $R_3 = -I(3)$, $\lambda^*(3)=\lambda_3$ and $\widehat{I}(\lambda^*(3)) = \ln(M(\lambda_3))$, thus we must have $-R_3 = 3\lambda_3 - \ln(M(\lambda_3))$. And indeed this is consistent with (a),(b),(c) because we have $2\ln(2)-1 = 3\ln(2) - (\ln(2)+1)$.

2. (7 points) Let $Y_1, Y_2, \dots$ denote independent and identically distributed random variables with optimistic GEO(1/2) distribution. Let $M_n = \max\{Y_1,\dots,Y_n\}$. For some $c\in\mathbb{R}_+$ let $Z(n) := M_n - c\cdot\ln(n)$. Let $n_k := 2^k, k=0,1,2,\dots$.

   (a) How to choose the constant c if we want $Z(n_k)$ to converge in distribution as $k\to\infty$? What is the c.d.f. of the limiting distribution?

   (b) Does $Z(n)$ converge in distribution as $n\to\infty$ with the above choice of c? Why?

   Solution: Y_1 is the number of coins that you need to toss until you see the first heads. $Y_1 > k$ if and only if the first k coin tosses are all tails, thus $\mathbb{P}(Y_1 > k) = 2^{-k}, k=0,1,2,\dots$. Thus $\mathbb{P}(Y_1\leq k) = 1-2^{-k}, k=0,1,2,\dots$, thus the c.d.f. of the GEO(1/2) distribution is $F(x) = 1-2^{-\lfloor x\rfloor}$ if $x\geq 0$. The c.d.f. of M_n is $(1-2^{-\lfloor x\rfloor})^n \vee 0$. The c.d.f. of $Z(n)$ is
   $$G_n(x) := \mathbb{P}(Z(n)\leq x) = \mathbb{P}(M_n \leq c\cdot\ln(n)+x) = (1-2^{-\lfloor c\cdot\ln(n)+x\rfloor})^n \vee 0.$$
   Now if we fix $x\in\mathbb{R}$ then
   $$2^{-\lfloor c\cdot\ln(n)+x\rfloor} \asymp 2^{-c\cdot\ln(n)} = e^{\ln(2)(-c\cdot\ln(n))} = e^{\ln(n)\cdot(-c\ln(2))} = n^{-c\ln(2)}$$
   Remember that if $a_n\to 0$ and $a_nb_n\to\gamma$ then $(1-a_n)^{b_n}\to e^{-\gamma}$. In our case $a_n = 2^{-\lfloor c\cdot\ln(n)+x\rfloor}$ and $b_n = n$, so if we want $G_n(x)$ (or at least $G_{n_k}(x)$) to converge to a number that is strictly between 0 and 1 then we need $0<\gamma<\infty$. Now $a_nb_n \asymp n^{-c\ln(2)}\cdot n$, thus we must choose $c = \frac{1}{\ln(2)}$, because otherwise a_nb_n will go to either zero or infinity.

   (a) Now $c\cdot\ln(n_k) = \frac{1}{\ln(2)}\cdot\ln(2^k) = \frac{1}{\ln(2)}\cdot k\cdot\ln(2) = k$, thus for any $x\in\mathbb{R}$, for large enough n we have $G_{n_k}(x) = (1-2^{-\lfloor k+x\rfloor})^{2^k} = (1-\frac{2^{-\lfloor x\rfloor}}{2^k})^{2^k}$, thus $G_{n_k}(x)\to G(x) := \exp\left(2^{-\lfloor x\rfloor}\right)$ as $k\to\infty$, thus $Z(n_k)$ indeed converges in distribution to a random variable Z with c.d.f. $G(x)$ as $k\to\infty$.

   (b) The answer is no, as we now explain. Note that $Z(n_k) = M_{n_k}-k$ is an integer-valued random variable for each k and thus Z is also an integer-valued random variable. But if we choose $n_k' := \lfloor\sqrt{2}\cdot 2^k\rfloor$ then $c\cdot\ln(n_k')-k\to\frac{1}{2}$, thus by a similar calculation as above, $Z(n_k')$ will converge in distribution to a random variable whose distribution is concentrated on values that are halfway between consecutive integers. Since the weak limits along the subsequences n_k and n_k' are different, the whole sequence $Z(n)$ does not converge in distribution as $n\to\infty$.