Stochastic Optimisation - Problem Sheet 1

Dom Hutchinson

November 12, 2020

Question 1.
Let $X_1 \sim \text{Poisson}(X_1)$ and $X_2 \sim \text{Poisson}(X_2)$ be independent random variables. Using generation functions, show that $(X_1 + X_2)$ is a Poisson random variable with mean $(\lambda_1 + \lambda_2)$

Answer 1.
Let $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$ be independent random variables. The moment generating function for these random variables is

$$M_{X_i}(t) := \mathbb{E}[e^{X_i}t] = e^{\lambda_i(e^t-1)}$$

Now consider the moment generating function for

$$(X_1 + X_2)$$

$$\begin{array}{rcl} M_{X_1+X_2}(t) & := & \mathbb{E}[e^{(X_1+X_2)t}] \\ & = & \mathbb{E}[e^{X_1t}e^{X_2t}] \\ & = & \mathbb{E}[e^{X_1t}]\mathbb{E}[e^{X_2t}] \text{ by independence} \\ & = & e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)} \\ & = & e^{(\lambda_1\lambda_2)(e^t-1)} \end{array}$$

This is equivalent to the moment generating function of a $\text{Poisson}(\lambda_1 + \lambda_2)$ distribution. Thus

$$(X_1 + X_2) \sim \text{Poisson}(\lambda_1 + \lambda_2)$$

Question 2 (a) i)
Let $T \sim \text{Exp}(\mu)$. Show that $\mathbb{E}[T] = \frac{1}{\mu}]$ and that μT has an exponential distribution with parameter 1.

Answer 2 (a) i)
First I derive the expected value of T

$$\begin{array}{rcl} \mathbb{E}[T] & := & \displaystyle\int_{-\infty}^{\infty} x f_X(x)dx \\ & = & \displaystyle\int_{-\infty}^{0} x \cdot \mu e^{-\mu x}dx \\ & = & [-xe^{-\mu x}]_0^\infty + \displaystyle\int_0^\infty e^{-\mu x}dx \text{ by integration by parts} \\ & = & 0 + \left[-\frac{1}{\mu}e^{-\mu x}\right]_0^\infty \\ & = & 0 + \left[0 - \left(-\frac{1}{\mu}\right)\right] \\ & = & \frac{1}{\mu} \end{array}$$

Now I derive the distribution of (μT). Note the moment generating function of T is

$$M_T(t) := \mathbb{E}[e^{tT}] = \frac{\mu}{\mu - t}$$

Now consider the moment generating function of (μT).

$$\begin{array}{rcl} M_{\mu T}(t) & := & \mathbb{E}[e^{t\mu T}] \\ & = & \dfrac{\mu}{\mu - t\mu} \\ & = & \dfrac{1}{1-t} \end{array}$$

This is equivalent to the moment generating function of a Exponential(1) distribution. Thus

$$(\mu T) \sim \text{Exponential}(1)$$

Question 2 (a) ii)
Let $T \sim \text{Exp}(\mu)$. Show that T is *memoryless*. i.e.

$$\forall\, t, u > - \quad \mathbb{P}(T > t + u | T > u) = \mathbb{P}(T > t)$$

Answer 2 (a) ii)

$$\begin{array}{rcl} \mathbb{P}(T > t + u | T > u) & = & \dfrac{\mathbb{P}(T > t + u, t > u)}{\mathbb{P}(T > u)} \\ & = & \dfrac{\mathbb{P}(T > t + u)}{\mathbb{P}(T > u)} \text{ since } \{T > t + u\} \subset \{T > u\} \\ & = & \dfrac{e^{-\mu(t+u)}}{e^{-\mu u}} \\ & = & e^{-\mu t} \\ & = & \mathbb{P}(T > t) \end{array}$$

Question 2 (b) i)
Let $T_1 \sim \text{Exp}(\lambda_1)$ and $T_2 \sim \text{Exp}(\lambda_2)$ be independent random variables and define $T := \min\{T_1, T_2\}$.
Show that T is an exponential random variable with parameter $(\lambda_1 + \lambda_2)$

Answer 2 (b) i)

$$\begin{array}{rcl} \mathbb{P}(T \leq t) & = & \mathbb{P}(\min\{T_1, T_2\} \leq t) \\ & = & 1 - \mathbb{P}(\min\{T_1, T_2) > t \\ & = & 1 - \mathbb{P}(T_1 > t, T_2 > t) \\ & = & 1 - \mathbb{P}(T_1 > t)\mathbb{P}(T_2 > t) \text{ by independence} \\ & = & 1 - e^{-\lambda_1 t} e^{-\lambda_2 t} \\ & = & 1 - e^{-(\lambda_1 + \lambda_2)t} \end{array}$$

This is equivalent to the CDF of a Exponential$(\lambda_1 + \lambda_2)$ distribution. Thus

$$T \sim \text{Exp}(\lambda_1 + \lambda_2)$$

Question 2 (b) ii)
Let $T_1 \sim \text{Exp}(\lambda_1)$ and $T_2 \sim \text{Exp}(\lambda_2)$ be independent random variables and define $T := \min\{T_1, T_2\}$.

Show that the probability that $T = T_1$ is $\dfrac{\lambda_1}{\lambda_1+\lambda_2}$ and that this is independent of the value of T.

Answer 2 (b) ii)

$$\begin{array}{rcl}
\mathbb{P}(T = T_1) & = & \mathbb{P}(\min\{T_1, T_2\} = T_1) \\
& = & \mathbb{P}(T_1 \leq T_2) \\
& = & \displaystyle\int_0^\infty \int_0^t f_{T_1,T_2}(s,t)ds\ dt \\
& = & \displaystyle\int_0^\infty f_{T_2}(t) \int_0^t f_{T_1}(s)ds\ dt \text{ by independence} \\
& = & \displaystyle\int_0^\infty \lambda_2 e^{-\lambda_2 t} \int_0^t \lambda_1 e^{-\lambda_1 s} ds\ dt \\
& = & \displaystyle\int_0^\infty \lambda_2 e^{-\lambda_2 t} [e^{-\lambda_1 s}]_0^t dt \\
& = & \displaystyle\int_0^\infty \lambda_2 e^{-\lambda_2 t} [1 - e^{-\lambda_1 t}] dt \\
& = & \lambda_2 \displaystyle\int_0^\infty e^{-\lambda_2 t} [1 - e^{-\lambda_1 t}] dt \\
& = & \lambda_2 \left(\displaystyle\int_0^\infty e^{-\lambda_2 t} dt - \int_0^\infty e^{-(\lambda_1+\lambda_2)t} dt \right) \\
& = & \lambda_2 \left(\left[-\dfrac{1}{\lambda_2} e^{-\lambda_2 t} \right]_0^\infty - \left[-\dfrac{1}{\lambda_1+\lambda_2} e^{-(\lambda_1+\lambda_2)t} \right]_0^\infty \right) \\
& = & \lambda_2 \left(\left[0 + \dfrac{1}{\lambda_2} \right] - \left[0 + \dfrac{1}{\lambda_1+\lambda_2} \right] \right) \\
& = & 1 - \dfrac{\lambda_2}{\lambda_1+\lambda_2} \\
& = & \dfrac{\lambda_1}{\lambda_1+\lambda_2}
\end{array}$$

Question 3.
Let $X_1, X_2, \dots$ by iid random variables with distribution $\text{Bern}(p)$ with $p \in [0,1]$. Let $q \in [0,1]$ with $q > p$. Show that

$$\mathbb{P}\left(\sum_{i=1}^n X_i\right) > nq) \leq \exp(-nK(q;p)) \quad \text{where} \quad K(q;p) = q \ln \frac{q}{p} + (1-q) \ln \frac{1-q}{1-p}$$

with $x \ln x$ defined to be 0 if $x = 0$.

Answer 3.
Let $X_1, X_2, \dots$ by iid random variables with distribution $\text{Bern}(p)$ with $p \in [0,1]$. Let $q \in [0,1]$ with $q > p$.
The moment generating function of X is $\mathcal{M}_X(t) := \mathbb{E}[e^{tX}] = (p + e^t + (1-p))$.

By applying *Chernoff Bounds* we have that

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^n X_i > nq\right) &\leq \inf_{\theta>0} e^{-nq\theta}(\mathbb{E}[e^{\theta X}])^n \\ &= \inf_{\theta>0} e^{-nq\theta}(pe^\theta + (1-p))^n\end{aligned}$$

Consider the natural log of the right hand side and define

$$f := -nq\theta + \ln(pe^\theta + 1 - p)$$

Since the natural log is a monotonically increasing function, $\inf\limits_{\theta>0} e^f$ is equal to the RHS of above. First I shall derive $\underset{\theta;\theta>0}{\text{argmin}}(f)$

$$\begin{aligned}\frac{\partial f}{\partial \theta} &= -nq + n\frac{pe^\theta}{pe^\theta+1-p} \\ \text{Setting} \qquad \frac{\partial f}{\partial \theta} &= 0 \\ \implies \qquad 0 &= -nq + n\frac{pe^\theta}{pe^\theta+1-p} \\ \implies \qquad q &= \frac{pe^\theta}{pe^\theta+1-p} \\ \implies \qquad pe^\theta + 1 - p &= \frac{p}{q}e^\theta \\ \implies \qquad e^\theta &= \frac{1-p}{\frac{p}{q}-p} \\ &= \frac{q-qp}{p-qp} \\ \implies \qquad \theta &= \ln\left(\frac{q-qp}{p-qp}\right)\end{aligned}$$

$$\begin{aligned}\text{Since} \qquad q &> p \\ \implies \qquad q-qp &> p-qp \\ \implies \qquad \frac{q-qp}{p-qp} &> 1 \\ \implies \qquad \ln\left(\frac{q-qp}{p-qp}\right) &> 0\end{aligned}$$

Thus $\underset{\theta;\theta>0}{\text{argmin}}(f) = \ln\left(\frac{q-qp}{p-qp}\right)$. This means

$$\begin{aligned}\inf_{\theta>0} f &= -nq\ln\left(\frac{q-qp}{p-qp}\right) + n\ln\left(p\cdot\frac{q-qp}{p-qp}+1-p\right) \\ &= -n\left[q\ln\left(\frac{q(1-p)}{p(1-q)}\right) - \ln\left(p\cdot\frac{q(1-p)}{p(1-q)}+1-p\right)\right] \\ &= -n\left[q\ln\left(\frac{q}{p}\right) + q\ln\left(\frac{1-p}{1-q}\right) - \ln\left(\frac{q(1-p)}{1-q}+1-p\right)\right] \\ &= -n\left[q\ln\left(\frac{q}{p}\right) - q\ln\left(\frac{1-q}{1-p}\right) - \ln\left(\frac{1-p}{1-q}\right)\right] \\ &= -n\left[q\ln\left(\frac{q}{p}\right) + (1-q)\ln\left(\frac{1-q}{1-p}\right)\right] \\ &= -nK(q;p) \\ \implies \qquad \inf_{\theta>0} e^{-nq\theta}(pe^\theta+(1-p))^n &= \exp(-nK(q;p)) \\ \implies \qquad \mathbb{P}\left(\sum_{i=1}^n X_i > nq\right) &\leq \exp(-nK(q;p))\end{aligned}$$

Question 4.
Let $X_1, X_2, \dots$ be IID Poisson random variables with parameter $\lambda \in \mathbb{R}^+$. Suppose $\mu \in \mathbb{R}^+$ is greater than λ. Show that

$$\mathbb{P}\left(\sum_{i=1}^n X_i < n\mu\right) \leq \exp(-nI(x;\lambda)) \quad \text{where} \quad I(\mu;\lambda) := \mu \ln\left(\frac{\mu}{\lambda}\right) - \mu + \lambda$$

Answer 4.
Let $X_1, X_2, \dots$ be IID Poisson random variables with parameter $\lambda \in \mathbb{R}^+$. Suppose $\mu \in \mathbb{R}^+$ is greater than λ.
Note that the moment generating function of each X_i is

$$M_{X_i}(\theta) := \mathbb{E}[e^{\theta X}] = e^{\lambda(e^\theta - 1)}$$

Chernoff Bounds state that

$$\begin{array}{rcl}\mathbb{P}\left(\sum_{i=1}^n X_i < n\mu\right) & \leq & \inf_{\theta<0} e^{-n\mu\theta} M_X(\theta)^n \\ & = & \inf_{\theta<0} e^{-n\mu\theta} e^{n\lambda(e^\theta-1)} \\ & = & \inf_{\theta<0} e^{n[\lambda(e^\theta-1)-\mu\theta]}\end{array}$$

Consider the natural log of the right hand side and define

$$f := n[\lambda(e^\theta - 1) - \mu\theta]$$

Since the natural log is a monotonically increasing function, $\inf_{\theta>0} e^f$ is equal to the RHS of above.
First I shall derive $\underset{\theta;\theta<0}{\text{argmin}}(f)$

$$\begin{array}{rrcl} & \frac{\partial f}{\partial \theta} & = & n(\lambda e^\theta - \mu) \\ \text{Setting} & \frac{\partial f}{\partial \theta} & = & 0 \\ \implies & n(\lambda e^\theta - \mu) & = & 0 \\ \implies & e^\theta & = & \frac{\mu}{\lambda} \\ \implies & \theta & = & \ln\left(\frac{\mu}{\lambda}\right) \\ & & & \\ \text{Since} & \mu & < & \lambda \\ \implies & \frac{\mu}{\lambda} & < & 0 \\ \implies & \ln\left(\frac{\mu}{\lambda}\right) & > & 0 \end{array}$$

Thus $\underset{\theta;\theta<0}{\text{argmin}}(f) = \ln\left(\frac{\mu}{\lambda}\right)$. This means

$$\begin{array}{rrcl} & \inf_{\theta<0} n[\lambda(e^\theta - 1) - \mu\theta] & = & n\left(\lambda\left(\frac{\mu}{\lambda} - 1\right) - \mu\ln\left(\frac{\mu}{\lambda}\right)\right) \\ & & = & n(\mu - \lambda) - n\mu\ln\left(\frac{\mu}{\lambda}\right) \\ & & = & -n(\mu\ln\left(\frac{\mu}{\lambda}\right) + \lambda - \mu) \\ & & = & -nI(\mu;\lambda) \\ \implies & \inf_{\theta<0} e^{n[\lambda(e^\theta-1)-\mu\theta]} & = & \exp(-nI(\mu;\lambda)) \\ \implies & \mathbb{P}\left(\sum_{i=1}^n X_i < n\mu\right) & \leq & \exp(-nI(\mu;\lambda)) \end{array}$$

Question 5.
Let $X_1, X_2, \dots$ be IID exponential random variables with parameter $\lambda \in \mathbb{R}^+$. Suppose $x \in \mathbb{R}^+$ is greater than $\frac{1}{\lambda}$. Show that

$$\mathbb{P}\left(\sum_{i=1}^n X_i > nx\right) \leq \exp(-nJ(x;\lambda)) \quad \text{where} \quad J(x;\lambda) := \lambda x - 1 - \ln(\lambda x)$$

Answer 5.
Let $X_1, X_2, \dots$ be IID exponential random variables with parameter $\lambda \in \mathbb{R}^+$. Suppose $x \in \mathbb{R}^+$ is greater than $\frac{1}{\lambda}$.
Note that the moment generating function of each X_i is

$$M_{X_i}(\theta) := \mathbb{E}[e^{\theta X_i}] = \frac{\lambda}{\lambda - \theta} \text{ for } \theta < \lambda$$

Chernoff Bounds state that

$$\begin{array}{rcl} \mathbb{P}\left(\sum_{i=1}^n X_i > nx\right) & \leq & \inf_{\theta>0} e^{-nx\theta} M_X(\theta)^n \\ & = & \inf_{\theta>0} e^{-nx\theta}\left(\frac{\lambda}{\lambda-\theta}\right)^n \end{array}$$

Consider the natural log of the right hand side and define

$$f := -nx\theta + n[\ln(\lambda) - \ln(\lambda - \theta)]$$

Since the natural log is a monotonically increasing function, $\inf_{\theta>0} e^f$ is equal to the RHS of above.

First I shall derive $\underset{\theta;\theta<0}{\text{argmin}}(f)$

$$\begin{array}{rrcl} & \frac{\partial f}{\partial \theta} & = & -nx+n\left[0-\frac{-1}{\lambda-\theta}\right] \\ & & = & -nx+\frac{n}{\lambda-\theta} \\ \text{Setting} & \frac{\partial f}{\partial \theta} & = & 0 \\ \implies & n\left(-x+\frac{1}{\lambda-\theta}\right) & = & 0 \\ \implies & \frac{1}{x} & = & \lambda-\theta \\ \implies & \theta & = & \lambda-\frac{1}{x} \\ & & & \\ \text{Since} & x & > & \frac{1}{\lambda} \\ \implies & \lambda>\frac{1}{x} & & \\ \implies & \lambda-\frac{1}{x} & > & 0 \end{array}$$

Thus $\underset{\theta;\theta>0}{\text{argmin}} f=\lambda-\frac{1}{x}$. This means

$$\begin{array}{rrcl} & \inf_{\theta>0} f & = & -nx\left(\lambda-\frac{1}{x}\right)+n\left[\ln(\lambda)-\ln\left(\lambda-\lambda+\frac{1}{x}\right)\right] \\ & & = & -n\left[x\lambda-1-\left(\ln(\lambda)-\ln\left(\frac{1}{x}\right)\right)\right] \\ & & = & -n[n\lambda-1-\ln(x\lambda)] \\ & & = & -nJ(x;\lambda) \\ \implies & \underset{\theta>0}{e}^{-nx\theta\left(\frac{\lambda}{\lambda-\theta}\right)^n} & = & \exp(-nJ(x;\lambda)) \\ \implies & \mathbb{P}\left(\sum_{i=1}^n X_i>nx\right) & \leq & \exp(-nJ(x;\lambda)) \end{array}$$

Question 6a)
Let $Z\sim N(0,1)$. Show that the moment generating function of Z is given by $\mathbb{E}[e^{\theta Z}]=e^{\frac{1}{2}\theta^2}$.

Answer 6a)
Let $Z\sim N(0,1)$.

$$\begin{aligned}
\mathbb{E}[e^{\theta Z}] &= \int_{-\infty}^{\infty} e^{\theta x} f_Z(x)dx \\
&= \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2-2x\theta)} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2+\frac{1}{2}\theta^2} dx \\
&= e^{\frac{1}{2}\theta^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2} dx \\
&= e^{\frac{1}{2}\theta^2} \int_{-\infty}^{\infty} f_Y(x)dx \quad \text{where} \quad Y \sim N(\theta, 1) \\
&= e^{\frac{1}{2}\theta^2} \cdot 1 \\
&= e^{\frac{1}{2}\theta^2}
\end{aligned}$$

Question 6b)
Let $X_1, X_2, \dots$ be iid random variables with distribution $N(\mu, \sigma^2)$ for $\mu, \sigma \in \mathbb{R}$. Let $\gamma \in \mathbb{R}$ st $\gamma > \mu$. Show that

$$\mathbb{P}\left(\sum_{i=1}^{n} > n\gamma\right) \leq \exp\left(-n\frac{(\gamma-\mu)^2}{2\sigma^2}\right)$$

Answer 6b)
Let $X_1, X_2, \dots$ be iid random variables with distribution $N(\mu, \sigma^2)$ for $\mu, \sigma \in \mathbb{R}$. Let $\gamma \in \mathbb{R}$ st $\gamma > \mu$.
The moment generating function of X is $\mathcal{M}_X(t) := \mathbb{E}[e^{tX}] = e^{\mu\theta + \frac{1}{2}\sigma^2\theta^2}$.
By applying *Chernoff Bounds* we have that

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i > n\gamma\right) &\leq \inf_{\theta>0} e^{-n\gamma\theta}(\mathbb{E}[e^{X\theta}])^n \\
&= \inf_{\theta>0} e^{-n\theta(\gamma-\mu-\frac{1}{2}\sigma^2\theta)}
\end{aligned}$$

Consider the natural log of the right hand side and define $f := -n\theta(\gamma - \mu - \dfrac{1}{2}\sigma^2\theta)$.
Since the natural log is a monotonically increasing function, $\inf\limits_{\theta>0} e^f$ is equal to the RHS of above.
First I shall derive $\operatorname*{argmin}\limits_{\theta;\theta>0}(f)$

$$\begin{aligned}
\frac{\partial f}{\partial \theta} &= -n(\gamma - \mu - \frac{1}{2}\sigma^2\theta) - n\theta(-\frac{1}{2}\sigma^2) \\
&= -n(\gamma - \mu - \sigma^2\theta)
\end{aligned}$$

$$\begin{aligned}
\text{Setting} \quad \frac{\partial f}{\partial \theta} &= 0 \\
\implies \quad \gamma - \mu - \sigma^2\theta &= 0 \\
\implies \quad \theta &= \frac{\gamma-\mu}{\sigma^2}
\end{aligned}$$

$$\begin{aligned}
\text{Since} \quad \gamma > \mu \quad &\& \quad \sigma^2 > 0 \\
\implies \quad 0 &< \frac{\gamma-\mu}{\sigma^2} = \theta
\end{aligned}$$

Thus $\underset{\theta;\theta>0}{\text{argmin}}(f) = \dfrac{\gamma-\mu}{\sigma^2}$. This means

$$\begin{array}{rrcl} & \inf_{\theta>0} f &=& -n\left(\dfrac{\gamma-\mu}{\sigma^2}\right)\left(\gamma-\mu-\dfrac{1}{2}(\gamma-\mu)\right) \\ & &=& -n\dfrac{(\gamma-\mu)^2}{2\sigma^2} \\ \implies & \inf_{\theta>0} e^{-n\theta\left(\gamma-\mu-\frac{1}{2}\sigma^2\theta\right)} &=& \exp\left(-n\dfrac{(\gamma-\mu)^2}{2\sigma^2}\right) \\ \implies & \mathbb{P}\left(\sum_{i=1}^n X_i > n\gamma\right) &\leq& \exp\left(-n\dfrac{(\gamma-\mu)^2}{2\sigma^2}\right) \end{array}$$

Question 7.
Let $\mathbf{X}$ be a random vector in $\mathbb{R}^d$ with moment generating function

$$M_{\mathbf{X}}(\boldsymbol{\theta}) := \mathbb{E}[\exp(\langle\boldsymbol{\theta},\mathbf{X}\rangle)] = \mathbb{E}[exp(\boldsymbol{\theta}^T\mathbf{X})] \quad \text{for } \boldsymbol{\theta}\in\mathbb{R}^d$$

Let $H(\boldsymbol{\theta},y)$ denote the half-space in $\mathbb{R}^d$ given by $H(\boldsymbol{\theta},y) := \{\mathbf{x}\in\mathbb{R}^d : \langle\boldsymbol{\theta},\mathbf{x}\rangle \geq y\}$ Show that

$$\forall\ \eta>0 \quad \mathbb{P}(\mathbf{X}\in H(\boldsymbol{\theta},y)) \leq e^{-\eta y}M_{\mathbf{X}}(\eta\boldsymbol{\theta})$$

Answer 7.
Define $Y=\boldsymbol{\theta}^T\mathbf{X}$, thus $(\eta Y) = (\eta\boldsymbol{\theta})^T\mathbf{X}$. This means the following events are equivalent

$$\{\mathbf{X}\in H(\boldsymbol{\theta},y)\} \Leftrightarrow \{\boldsymbol{\theta}^T\mathbf{X}\geq y\} \Leftrightarrow \{\eta(\boldsymbol{\theta}^T\mathbf{X}) =: \eta Y \geq \eta y\}\ \forall\ \eta\geq 0$$

Thus

$$\begin{array}{rcl} \mathbb{P}(\mathbf{X}\in H(\boldsymbol{\theta},y)) &\leq& \mathbb{P}(\eta Y\geq \eta y) \\ &\leq& \mathbb{P}(e^{\eta Y}\geq e^{\eta y}) \\ &\leq& \dfrac{\mathbb{E}[\exp(\eta Y)]}{e^{\eta y}} \text{ by Markov's Inequality} \\ &=& e^{-\eta y}M_{\mathbf{X}}(\eta\boldsymbol{\theta}) \end{array}$$

Question 8.
Let $X_1,X_2,\dots$ be IID random variables with mean μ and taking values in the interval $[a,b]$ where a and b are finite. Show that

$$\mathbb{P}\left(\sum_{i=1}^n (X_i-\mu) > nt\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

Answer 8.
Let $X_1,X_2,\dots$ be IID random variables with mean μ and taking values in the interval $[a,b]$ where a and b are finite.
Define random variable $Y_i := \dfrac{X_i-a}{b-a}$. Note that $Y_i\in[0,1]$ and Y_i has mean $\mu_Y := \dfrac{\mu-a}{b-a}$. By

Hoeffding's Inequality

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^n (Y_i - \mu_Y)?ns\right) &\leq \exp(-2ns^2)\ \forall\ s > 0 \\
\implies \mathbb{P}\left(\sum_{i=1}^n \left(\frac{X_i - a}{b-a} - \frac{\mu - a}{b-a}\right) > ns\right) &\leq \exp(-2ns^2) \\
\implies \mathbb{P}\left(\frac{1}{b-a}\sum_{i=1}^n (X_i \mu) > ns\right) &\leq \exp(-2ns^2) \\
\implies \mathbb{P}\left(\sum_{i=1}^n (X_i \mu) > (b-a)ns\right) &\leq \exp(-2ns^2) \text{ since } b > a
\end{aligned}$$

By defining $t := s(b-a)$, and noting that $s = \dfrac{t}{b-a}$, we get the desired result

$$\mathbb{P}\left(\sum_{i=1}^n (X_i) - n\mu > tn\right) \leq \exp\left(-2n\frac{t^2}{(b-a)^2}\right)$$