Problem 4.

It took around 7 hours for me to complete the homework.

Problem 1.

Let X follows the exponential distribution with the rate parameter λ, then its (uncentered) MGF is

$$\mathbb{E}\left[\exp\left(tX\right)\right] = \frac{\lambda}{\lambda - t}, \ \forall t < \lambda$$

and its mean is

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

With the same r.v. X, we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(t(X - \mathbb{E}[X])\right)\right] \\
&= \frac{1}{\exp\left(t\,\mathbb{E}[X]\right)}\,\mathbb{E}\left[\exp\left(tX\right)\right] \\
&= \frac{1}{\exp\left(\frac{t}{\lambda}\right)}\,\mathbb{E}\left[\exp\left(tX\right)\right] \\
&= \frac{1}{\exp\left(\frac{t}{\lambda}\right)}\frac{\lambda}{\lambda - t} \\
&= \frac{1}{\exp\left(\frac{t}{\lambda}\right)}\left(1 + \frac{t}{\lambda - t}\right) \\
&\le \frac{1}{\exp\left(\frac{t}{\lambda}\right)}\exp\left(\frac{t}{\lambda - t}\right) && \text{(bc } 1 + x \le \exp\left(x\right), \ \forall x\text{)} \\
&= \exp\left(\frac{t}{\lambda - t} - \frac{t}{\lambda}\right) \\
&= \exp\left(\frac{t^2}{\lambda(\lambda - t)}\right) \\
&= \exp\left(\frac{t^2\frac{2}{\lambda^2}}{2(1 - \frac{1}{\lambda}t)}\right) \\
&\le \exp\left(\frac{t^2\frac{2}{\lambda^2}}{2(1 - \frac{1}{\lambda}|t|)}\right), \ \forall |t| \le \lambda && \text{(from the assumption of exponential distribution } t \le \lambda\text{)}
\end{aligned}$$

From the definition of Sub Gamma distribution, we want to choose the smallest parameters which are $c = \frac{1}{\lambda}$ and $\nu^2 = \frac{2}{\lambda^2}$

Problem 3.

Problem 3.a.

By definition, we have

$$Q = (1-2\varepsilon)\begin{pmatrix} 1 & -\tau \\ -\tau & \tau^2 \end{pmatrix} + \varepsilon \begin{pmatrix} 1 & \tau \\ \tau & \tau^2 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

We set $\varepsilon = \tau^2$, then

$$\begin{aligned} Q &= (1-2\tau^2)\begin{pmatrix} 1 & -\tau \\ -\tau & \tau^2 \end{pmatrix} + \tau^2 \begin{pmatrix} 1 & \tau \\ \tau & \tau^2 \end{pmatrix} + \tau^2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1-\tau^2 & 3\tau^3 - r \\ 3\tau^3 - r & -\tau^4 + 2\tau^2 \end{pmatrix} \end{aligned}$$

Use wolfram alpha to find the Q^{-1}, we have

$$Q^{-1} = \frac{1}{\tau^2(8\tau^4 - 3\tau^2 - 1)} \begin{pmatrix} \tau^2(\tau^2-2) & r(3\tau^3-1) \\ r(3\tau^3-1) & \tau^2 - 1 \end{pmatrix}$$

Problem 3.b.

From Q^{-1}, we have $(Q^{-1})_{11} = \frac{\tau^4 - 2\tau^2}{\tau^2(8\tau^4-3\tau^2-1)}$

Compute R_1, R_2, and R_3

$$\begin{aligned} R_1 &= \left| e_1^T Q^{-1} x^{(1)} \right| \\ &= \left| \frac{-2\tau^4 - \tau^2}{\tau^2(8\tau^4 - 3\tau^2 - 1)} \right| \\ &= \left| \frac{2\tau^4 + \tau^2}{\tau^2(8\tau^4 - 3\tau^2 - 1)} \right| \end{aligned}$$

$$\begin{aligned} R_2 &= \left| e_1^T Q^{-1} x^{(2)} \right| \\ &= \left| \frac{4\tau^4 - 3\tau^2}{\tau^2(8\tau^4 - 3\tau^2 - 1)} \right| \end{aligned}$$

$$\begin{aligned} R_3 &= \left| e_1^T Q^{-1} x^{(3)} \right| \\ &= \left| \frac{3\tau^3 - \tau}{\tau^2(8\tau^4 - 3\tau^2 - 1)} \right| \end{aligned}$$

Because $\tau \le 1$, the term order of τ will dominate other term, hence orderwise, R_3 is the largest range. And orderwise, R_3 can be much larger than $(Q^{-1})_{11}$. So $i^* = 3$

Compute the $\frac{R_3}{(Q^{-1})_{11}}$

$$\frac{R_3}{(Q^{-1})_{11}} = \left| \frac{3\tau^3 - \tau}{\tau^4 - 2\tau^2} \right| = O\left(\frac{1}{\tau}\right)$$

So the ratio $\frac{R_3}{(Q^{-1})_{11}}$ scales with $O\left(\frac{1}{\tau}\right)$ with $\tau < 1$. The case shows that the range R_3 can be much larger $(Q^{-1})_{11}$, and hence the Catoni estimator step of PopArt is necessary to perform well in the nonasymptotic regime.