

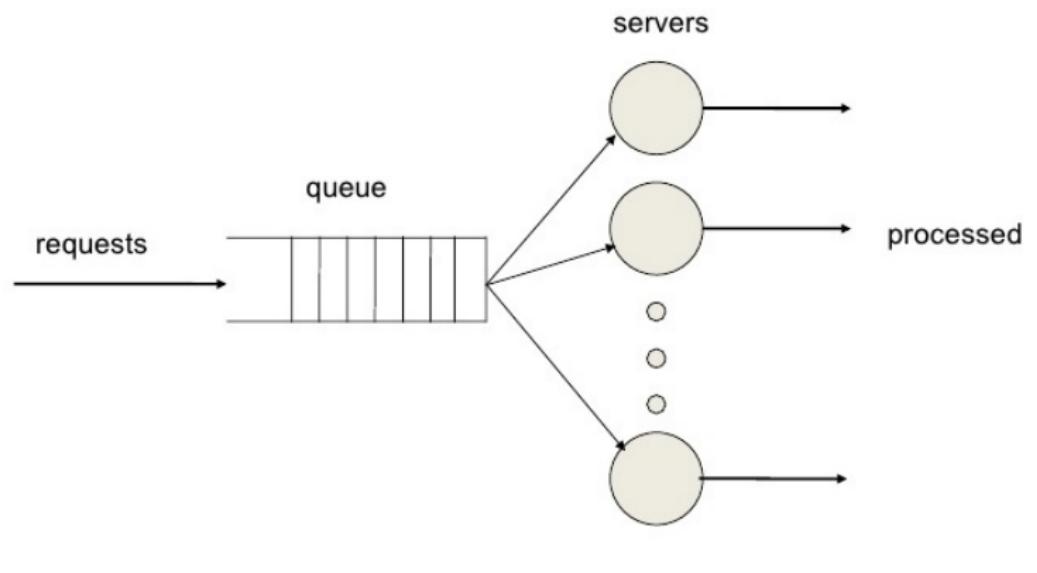
On deriving exact asymptotics for applications in queuing

Julia Kuhn

*The University of Queensland,
University of Amsterdam*

10 January 2017

A queuing system



Example: Call center

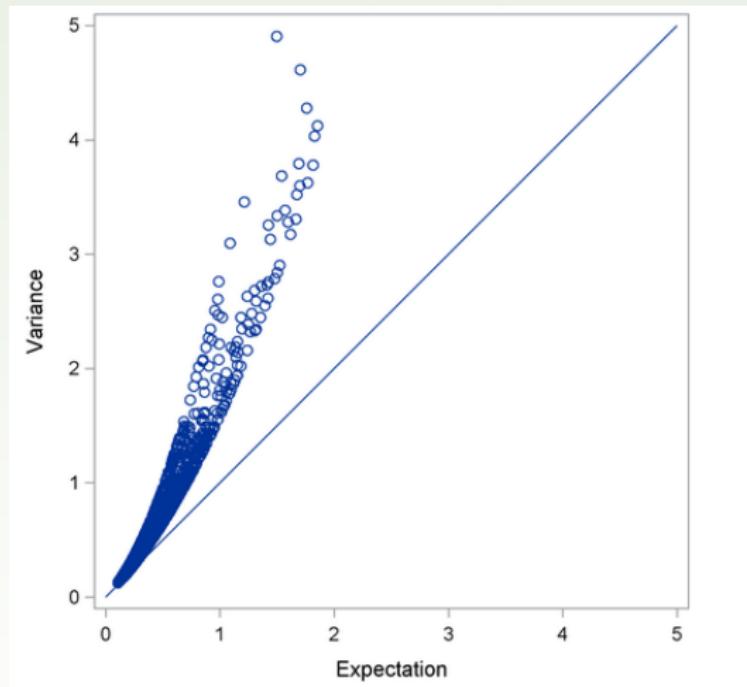


Usual assumptions

- arrivals are Poisson with constant rate λ
- service durations are exponential with constant rate μ

Does call center data really look like this?

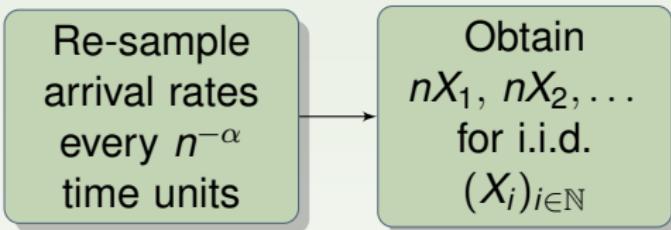
Arrivals are overdispersed



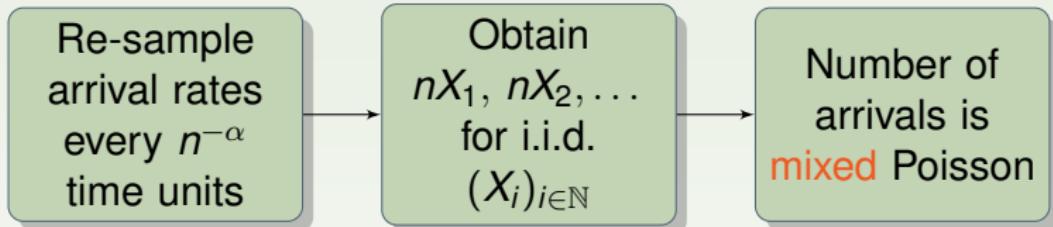
An overdispersed arrival process

Re-sample
arrival rates
every $n^{-\alpha}$
time units

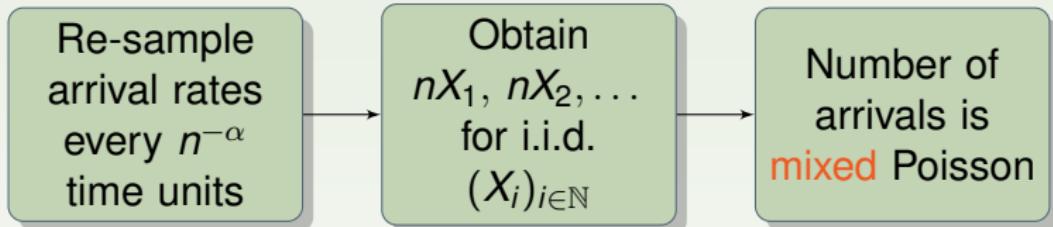
An overdispersed arrival process



An overdispersed arrival process



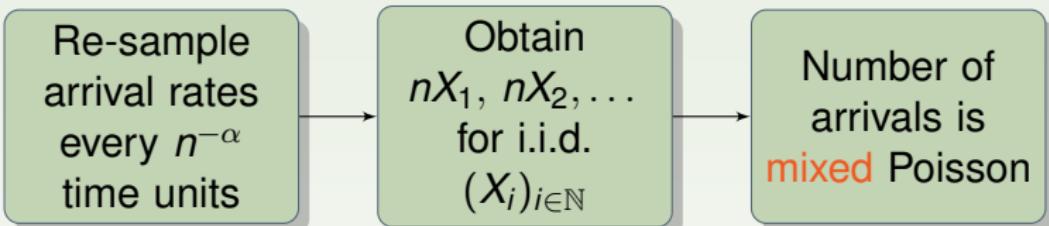
An overdispersed arrival process



Total number of arrivals at time t is Poisson with random mean-value function

$$\int_0^t r(s)ds = n\bar{X}_{n^\alpha} := n^{1-\alpha} \sum_{i=1}^{n^\alpha t} X_i.$$

An overdispersed arrival process



Total number of arrivals at time t is Poisson with random mean-value function

$$\int_0^t r(s)ds = n\bar{X}_{n^\alpha} := n^{1-\alpha} \sum_{i=1}^{n^\alpha t} X_i.$$

What is $P_n(a) := \mathbb{P} \left(\text{Pois} \left(n\bar{X}_{n^\alpha} \right) \geq na \right)$ for $a > \mathbb{E}X_i$?

Another example

What about

$$\mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{n,i} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{n,i} \right)$$

where $\mathbb{E} X > \mathbb{E} Y$?

Another example

What about

$$\mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{n,i} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{n,i} \right)$$

where $\mathbb{E} X > \mathbb{E} Y$?

Applications:

- $\mathbb{P} \left(\min_{i \in \{1, \dots, d_X\}} \bar{X}_{n,i} \leq \max_{i \in \{1, \dots, d_Y\}} \bar{Y}_{n,i} \right)$, the probability of false classification in LLR testing for anomaly identification

Another example

What about

$$\mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{n,i} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{n,i} \right)$$

where $\mathbb{E} X > \mathbb{E} Y$?

Applications:

- $\mathbb{P} \left(\min_{i \in \{1, \dots, d_X\}} \bar{X}_{n,i} \leq \max_{i \in \{1, \dots, d_Y\}} \bar{Y}_{n,i} \right)$, the probability of false classification in LLR testing for anomaly identification
- $\mathbb{P} \left(\exists i \in \{1, \dots, d - k + 1\} : \bar{X}_{n,(i)} \leq \bar{Y}_{n,(k+i-1)} \right)$ with applications in packing or queuing problems

Applications of Mathematics

38

Amir Dembo
Ofer Zeitouni

**Large Deviations
Techniques
and Applications**

Second Edition

 Springer

Theorem (Cramér)

Let $(Y_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables taking values in \mathbb{R} and with finite MGF $M(\theta)$. Then for $a > \mathbb{E} Y_i$ the sequence of sample means \bar{Y}_n satisfies

$$\frac{1}{n} \log \mathbb{P} (\bar{Y}_n \geq a) \rightarrow -\inf_{x \geq a} I(x)$$

as $n \rightarrow \infty$, where

$$I(a) := \sup_{\theta} \{\theta a - \log M(\theta)\}$$

is the Legendre transform of $\log M(\theta)$.

Theorem (Cramér)

Let $(Y_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables taking values in \mathbb{R} and with finite MGF $M(\theta)$. Then the sequence of sample means \bar{Y}_n satisfies

$$\frac{1}{n} \log \mathbb{P} (\bar{Y}_n \geq a) \rightarrow -I(a)$$

as $n \rightarrow \infty$, where

$$I(a) := \sup_{\theta} \{\theta a - \log M(\theta)\}$$

is the Legendre transform of $\log M(\theta)$.

How accurate is this?

$$\frac{1}{n} \log \mathbb{P} (\bar{Y}_n \geq a) \rightarrow -I(a)$$

This suggests the approximation

$$\mathbb{P} (\bar{Y}_n \geq a) \approx e^{-nI(a)}.$$

How accurate is this?

$$\frac{1}{n} \log \mathbb{P} (\bar{Y}_n \geq a) \rightarrow -I(a)$$

This suggests the approximation

$$\mathbb{P} (\bar{Y}_n \geq a) \approx e^{-nI(a)}.$$

But note that Cramér actually only says that

$$\mathbb{P} (\bar{Y}_n \geq a) = \xi(n) e^{-nI(a)}, \quad \frac{1}{n} \log \xi(n) \rightarrow 0.$$

How accurate is this?

$$\frac{1}{n} \log \mathbb{P} (\bar{Y}_n \geq a) \rightarrow -I(a)$$

This suggests the approximation

$$\mathbb{P} (\bar{Y}_n \geq a) \approx e^{-nI(a)}.$$

But note that Cramér actually only says that

$$\mathbb{P} (\bar{Y}_n \geq a) = \xi(n) e^{-nI(a)}, \quad \frac{1}{n} \log \xi(n) \rightarrow 0.$$

Thus, $\xi(n) = 10^{10}$, or n^{100} , or even $\exp(n^{0.99})$ are possible!

Identify $\xi(n) \Rightarrow$ Exact asymptotics

Theorem (Bahadur-Rao)

Assume the conditions of Cramér. Then, provided that $\theta(a)$ exists such that

$$I(a) := \sup_{\theta} \{\theta a - \Lambda(\theta)\} = \theta(a) a - \Lambda(\theta(a)),$$

we have

$$\mathbb{P}(\bar{Y}_n \geq a) \sim \frac{1}{\sqrt{n}} C(a) e^{-nI(a)}$$

where (for non-lattice Y_i) $C(a) = \left(\theta(a) \sqrt{2\pi \Lambda''(\theta(a))} \right)^{-1}$.

Sketch of proof

Idea: Apply a change of measure so that event of interest is not rare

⇒ Central limit arguments apply
(Berry-Esseen)

Sketch of proof

Chernoff bound:

$$\mathbb{P}(\bar{Y}_n \geq a) = \mathbb{P}\left(e^{\theta \sum_{i=1}^n Y_i} \geq e^{\theta n a}\right) \leq e^{-\theta n a} \mathbb{E}\left[e^{\theta \sum_{i=1}^n Y_i}\right]$$

Sketch of proof

Chernoff bound:

$$\mathbb{P}(\bar{Y}_n \geq a) = \mathbb{P}\left(e^{\theta \sum_{i=1}^n Y_i} \geq e^{\theta n a}\right) \leq e^{-\theta n a} \mathbb{E}\left[e^{\theta \sum_{i=1}^n Y_i}\right]$$

Therefore:

$$\frac{1}{n} \log \mathbb{P}\left(\bar{Y}_n \geq a\right) \leq -\max_{\theta} \{\theta a - \log M(\theta)\} = -I(a)$$

Sketch of proof

Chernoff bound:

$$\mathbb{P}(\bar{Y}_n \geq a) = \mathbb{P}\left(e^{\theta \sum_{i=1}^n Y_i} \geq e^{\theta n a}\right) \leq e^{-\theta n a} \mathbb{E}\left[e^{\theta \sum_{i=1}^n Y_i}\right]$$

Therefore:

$$\frac{1}{n} \log \mathbb{P}\left(\bar{Y}_n \geq a\right) \leq -\max_{\theta} \{\theta a - \log M(\theta)\} = -I(a)$$

Note: Maximising θ^* satisfies $a - M'(\theta)/M(\theta) = 0$

Sketch of proof

Chernoff bound:

$$\mathbb{P}(\bar{Y}_n \geq a) = \mathbb{P}\left(e^{\theta \sum_{i=1}^n Y_i} \geq e^{\theta n a}\right) \leq e^{-\theta n a} \mathbb{E}\left[e^{\theta \sum_{i=1}^n Y_i}\right]$$

Therefore:

$$\frac{1}{n} \log \mathbb{P}\left(\bar{Y}_n \geq a\right) \leq -\max_{\theta} \{\theta a - \log M(\theta)\} = -I(a)$$

Note: Maximising θ^* satisfies $a - M'(\theta)/M(\theta) = 0$

⇒ To “make” \bar{Y}_n have mean a , define density $q(\cdot)$ s.t.:

$$\mathbb{E}_q[\bar{Y}_n] = \int_{-\infty}^{\infty} y q(y) dy = \frac{M'(\theta^*)}{M(\theta^*)} = \frac{\mathbb{E}_p[e^{\theta^* Y}]}{\mathbb{E}_p[e^{\theta^* Y}]}$$

Sketch of proof

Chernoff bound:

$$\mathbb{P}(\bar{Y}_n \geq a) = \mathbb{P}\left(e^{\theta \sum_{i=1}^n Y_i} \geq e^{\theta n a}\right) \leq e^{-\theta n a} \mathbb{E}\left[e^{\theta \sum_{i=1}^n Y_i}\right]$$

Therefore:

$$\frac{1}{n} \log \mathbb{P}\left(\bar{Y}_n \geq a\right) \leq -\max_{\theta} \{\theta a - \log M(\theta)\} = -I(a)$$

Note: Maximising θ^* satisfies $a - M'(\theta)/M(\theta) = 0$

⇒ To “make” \bar{Y}_n have mean a , define density $q(\cdot)$ s.t.:

$$\mathbb{E}_q[\bar{Y}_n] = \int_{-\infty}^{\infty} y q(y) dy = \frac{M'(\theta^*)}{M(\theta^*)} = \frac{\mathbb{E}_p[Y e^{\theta^* Y}]}{\mathbb{E}_p[e^{\theta^* Y}]}$$

⇒ Choose $q(y) = \frac{e^{\theta^* y}}{M(\theta^*)} p(y)$.

Can we apply Bahadur-Rao?

$$\mathbb{P} \left(\bar{Y}_n \geq a \right) \sim \frac{1}{\sqrt{n}} C(a) e^{-nl(a)}$$

Can we apply Bahadur-Rao?

$$\mathbb{P}(\bar{Y}_n \geq a) \sim \frac{1}{\sqrt{n}} C(a) e^{-nl(a)}$$

For $\alpha = 1$:

$$\mathbb{P}\left(\text{Pois}\left(n\bar{X}_{n^\alpha}\right) \geq na\right) = \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n \text{Pois}(X_i) \geq a\right)$$



Can we apply Bahadur-Rao?

$$\mathbb{P}(\bar{Y}_n \geq a) \sim \frac{1}{\sqrt{n}} C(a) e^{-nl(a)}$$

For $\alpha = 1$:

$$\mathbb{P}\left(\text{Pois}\left(n\bar{X}_{n^\alpha}\right) \geq na\right) = \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n \text{Pois}(X_i) \geq a\right)$$



For $\alpha \neq 1$:

?

A recipe for mixtures



- ① Identify point of concentration a^*
- ② Get exact asymptotics for lower bound $\int_{a^* - \varepsilon}^{a^* + \varepsilon}$, using, e.g.:
 - Bahadur-Rao
 - Taylor approximation
 - ...
- ③ Show that first and last summand of upper bound $\int_{-\infty}^{a^* - \varepsilon} + \int_{a^* - \varepsilon}^{a^* + \varepsilon} + \int_{a^* + \varepsilon}^{\infty}$ are negligible (and thus lower and upper bound coincide)

A guess for the rate function

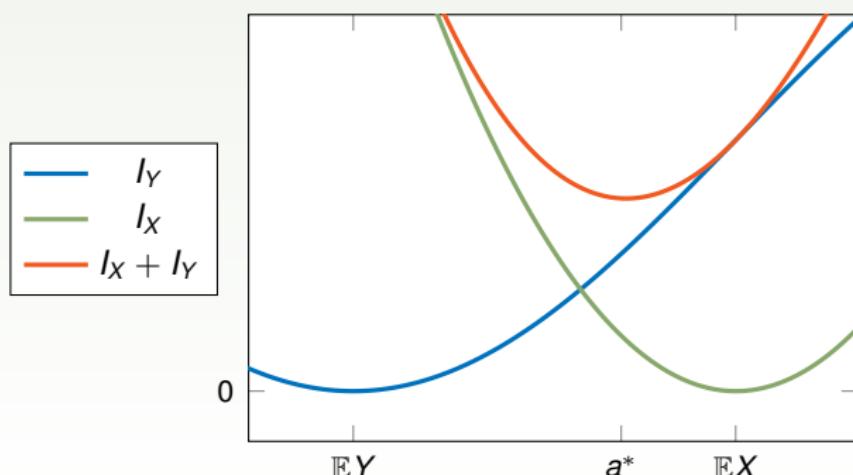
$$\begin{aligned}\alpha_{d_X, d_Y}(n) &:= \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \right) \\ &= \int_{-\infty}^{\infty} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq a \right) \mathbb{P} \left(\min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \in da \right)\end{aligned}$$

A guess for the rate function

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &:= \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \right) \\ &= \underbrace{\int_{-\infty}^{\infty} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq a \right)}_{\approx e^{-nd_X l_X(a)}} \mathbb{P} \left(\min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \in da \right) \\ &\quad \underbrace{\approx e^{-nd_Y l_Y(a)}}_{\approx e^{-nd_Y l_Y(a)}}\end{aligned}$$

A guess for the rate function

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &:= \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \right) \\ &= \underbrace{\int_{-\infty}^{\infty} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq a \right)}_{\approx e^{-nd_X I_X(a)}} \underbrace{\mathbb{P} \left(\min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \in da \right)}_{\approx e^{-nd_Y I_Y(a)}}\end{aligned}$$



$$\alpha_{d_X, d_Y}(n) := \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \leq \min_{i \in \{1, \dots, d_Y\}} \bar{Y}_{i,n} \right)$$

$$J(a) := d_X I_X(a) + d_Y I_Y(a)$$

$$a^* := \arg \min J(a)$$

$$K(a) := (-C_X(a))^{d_X} C_Y(a)^{d_Y} d_Y I'_Y(a)$$

Proposition (K., Mandjes, Taimre)

Suppose that X and Y are continuous, and fulfil the conditions of Bahadur-Rao. Then,

$$\alpha_{d_X, d_Y}(n) \sim n^{(1-d_X-d_Y)/2} K(a^*) \sqrt{\frac{2\pi}{J''(a^*)}} e^{-nJ(a^*)}.$$

Proof: Lower bound

Assume that that C_X , C_Y , K are smooth.

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &= \int_{-\infty}^{\infty} \left(\mathbb{P}(\bar{Y}_{1,n} \geq a) \right)^{d_Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) \\ &\gtrsim \int_{a^* - \varepsilon}^{a^* + \varepsilon} \left(\frac{C_Y(a)}{\sqrt{n}} e^{-nI_Y(a)} \right)^{d_Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) \\ &= \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots da + g(a^* + \varepsilon, n) - g(a^* - \varepsilon, n),\end{aligned}$$

where

$$g(a, n) := \left(\frac{C_Y(a)}{\sqrt{n}} e^{-nI_Y(a)} \right)^{d_Y} \left(-\frac{C_X(a)}{\sqrt{n}} e^{-nI_X(a)} \right)^{d_X}.$$

Proof: Lower bound

Assume that C_X, C_Y, K are smooth.

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &= \int_{-\infty}^{\infty} \left(\mathbb{P}(\bar{Y}_{1,n} \geq a) \right)^{d_Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) \\ &\gtrsim \int_{a^* - \varepsilon}^{a^* + \varepsilon} \left(\frac{C_Y(a)}{\sqrt{n}} e^{-nI_Y(a)} \right)^{d_Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) \\ &= \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots da + g(a^* + \varepsilon, n) - g(a^* - \varepsilon, n),\end{aligned}$$

where

$$g(a, n) := \left(\frac{C_Y(a)}{\sqrt{n}} e^{-nI_Y(a)} \right)^{d_Y} \left(-\frac{C_X(a)}{\sqrt{n}} e^{-nI_X(a)} \right)^{d_X}.$$

We will show that first term has decay rate $J(a^*)$
 $\Rightarrow g(a^* \pm \varepsilon, n)$ are negligible

Proof: Lower bound

So we focus on the remaining term:

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &\geq \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots e^{-nJ(a)} da \\ &= e^{-nJ(a^*)} \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots e^{-n[J(a) - J(a^*)]} da,\end{aligned}$$

Proof: Lower bound

So we focus on the remaining term:

$$\begin{aligned}\alpha_{d_X, d_Y}(n) &\geq \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots e^{-nJ(a)} da \\ &= e^{-nJ(a^*)} \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots e^{-n[J(a) - J(a^*)]} da,\end{aligned}$$

How does $e^{-n[J(a) - J(a^*)]}$ behave?

Proof: Lower bound

With $\psi(a) = o(a^2)$, by a Taylor expansion of $J(a)$ around a^* :

$$J(a) - J(a^*) \leq \frac{1}{2} J''(a^*)(a - a^*)^2 + \psi(a - a^*),$$

where we used that $J(a)$ is minimal at a^* .

Proof: Lower bound

With $\psi(a) = o(a^2)$, by a Taylor expansion of $J(a)$ around a^* :

$$J(a) - J(a^*) \leq \frac{1}{2} J''(a^*)(a - a^*)^2 + \psi(a - a^*),$$

where we used that $J(a)$ is minimal at a^* .

This upper bound yields

$$\int_{a^*-\varepsilon}^{a^*+\varepsilon} \dots e^{-n[J(a)-J(a^*)]} da \geq n^{1-(d_x+d_Y)/2} \inf_{a \in (a^*-\varepsilon, a^*+\varepsilon)} K(a) \int_{-\varepsilon}^{\varepsilon} e^{-n[\frac{1}{2}J''(a^*)a^2 + \psi(a)]} da$$

Proof: Lower bound

$$\int_{-\varepsilon}^{\varepsilon} e^{-n[\frac{1}{2}J''(a^*)a^2 + \psi(a)]} da$$

Apply transformation $b = \sqrt{nJ''(a^*)} a$:

$$\frac{1}{\sqrt{J''(a^*)}} \int_{-\varepsilon/\sqrt{nJ''(a^*)}}^{\varepsilon/\sqrt{nJ''(a^*)}} e^{-b^2/2 - n\psi(b/\sqrt{nJ''(a^*)})} db.$$

Proof: Lower bound

$$\int_{-\varepsilon}^{\varepsilon} e^{-n[\frac{1}{2}J''(a^*)a^2 + \psi(a)]} da$$

Apply transformation $b = \sqrt{nJ''(a^*)} a$:

$$\frac{1}{\sqrt{J''(a^*)}} \int_{-\varepsilon\sqrt{nJ''(a^*)}}^{\varepsilon\sqrt{nJ''(a^*)}} e^{-b^2/2 - n\psi(b/\sqrt{nJ''(a^*)})} db.$$

By dominated convergence, and since $\psi(a) = o(a^2)$:

$$\int_{-\varepsilon\sqrt{nJ''(a^*)}}^{\varepsilon\sqrt{nJ''(a^*)}} e^{-b^2/2 - n\psi(b/\sqrt{nJ''(a^*)})} db \rightarrow \int_{-\infty}^{\infty} e^{-b^2/2} db = \sqrt{2\pi}.$$

Proof: Lower bound

Letting $\varepsilon \rightarrow 0$, we have found:

$$\begin{aligned} & \liminf_{n \rightarrow \infty} n^{(d_X + d_Y - 1)/2} e^{nJ(a^*)} \alpha_{d_X, d_Y}(n) \\ & \geq K(a^*) \sqrt{\frac{2\pi}{J''(a^*)}}. \end{aligned}$$

Proof: Upper bound

Similarly, we can show:

$$\begin{aligned} \limsup_{n \rightarrow \infty} n^{(d_X + d_Y - 1)/2} e^{nJ(a^*)} \int_{a^* - \varepsilon}^{a^* + \varepsilon} \dots da \\ \geq K(a^*) \sqrt{\frac{2\pi}{J''(a^*)}}. \end{aligned}$$

⇒ It remains to be checked that $\int_{-\infty}^{a^* - \varepsilon} \dots da$ and $\int_{a^* + \varepsilon}^{\infty} \dots da$ are asymptotically negligible.

Proof: Upper bound

By a Chernoff bound,

$$\begin{aligned} & \int_{-\infty}^{a^* - \varepsilon} \mathbb{P}(\bar{Y}_{1,n} \geq a)^{d_Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \leq \int_{-\infty}^{\mathbb{E} Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \quad + \int_{\mathbb{E} Y}^{a^* - \varepsilon} e^{-nd_Y l_Y(a)} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \end{aligned}$$

Proof: Upper bound

By a Chernoff bound,

$$\begin{aligned} & \int_{-\infty}^{a^* - \varepsilon} \mathbb{P}(\bar{Y}_{1,n} \geq a)^{d_Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \leq \int_{-\infty}^{\mathbb{E} Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \quad + \int_{\mathbb{E} Y}^{a^* - \varepsilon} e^{-nd_Y l_Y(a)} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \end{aligned}$$

First term

$$\int_{-\infty}^{\mathbb{E} Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) = \mathbb{P} \left(\bar{X}_{i,n} \leq \mathbb{E} Y \right)^{d_X} \leq e^{-nd_X I_X(\mathbb{E} Y)}$$

First term

$$\int_{-\infty}^{\mathbb{E} Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) = \mathbb{P} \left(\bar{X}_{i,n} \leq \mathbb{E} Y \right)^{d_X} \leq e^{-nd_X I_X(\mathbb{E} Y)}$$

So the rate function is

$$-d_X I_X(\mathbb{E} Y) = - (d_X I_X(\mathbb{E} Y) + d_Y I_Y(\mathbb{E} Y))$$

First term

$$\int_{-\infty}^{\mathbb{E} Y} \mathbb{P} \left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da \right) = \mathbb{P} \left(\bar{X}_{i,n} \leq \mathbb{E} Y \right)^{d_X} \leq e^{-nd_X I_X(\mathbb{E} Y)}$$

So the rate function is

$$-d_X I_X(\mathbb{E} Y) = - (d_X I_X(\mathbb{E} Y) + d_Y I_Y(\mathbb{E} Y))$$

By definition of a^* this is smaller than

$$-(d_X I_X(a^*) + d_Y I_Y(a^*)) = -J(a^*)$$



Proof: Upper bound

By a Chernoff bound,

$$\begin{aligned} & \int_{-\infty}^{a^* - \varepsilon} \mathbb{P}(Y_{1,n} \geq a)^{d_Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \leq \int_{-\infty}^{\mathbb{E} Y} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \\ & \quad + \int_{\mathbb{E} Y}^{a^* - \varepsilon} e^{-nd_Y l_Y(a)} \mathbb{P}\left(\max_{i \in \{1, \dots, d_X\}} \bar{X}_{i,n} \in da\right) \end{aligned}$$

Proof: Upper bound

Integrating by parts we obtain an upper bound

$$\begin{aligned} & e^{-nd_Y l_Y(a^* - \varepsilon)} \mathbb{P} \left(\bar{X}_{1,n} \leq a^* - \varepsilon \right)^{d_X} \\ & + \int_{\mathbb{E} Y}^{a^* - \varepsilon} nd_Y l'_Y(a) e^{-nd_Y l_Y(a)} \mathbb{P} \left(\bar{X}_{1,n} \leq a \right)^{d_X} da \end{aligned}$$

Proof: Upper bound

Integrating by parts we obtain an upper bound

$$\begin{aligned} & e^{-nd_Y l_Y(a^* - \varepsilon)} \mathbb{P} \left(\bar{X}_{1,n} \leq a^* - \varepsilon \right)^{d_X} \\ & + \int_{\mathbb{E} Y}^{a^* - \varepsilon} nd_Y l'_Y(a) e^{-nd_Y l_Y(a)} \mathbb{P} \left(\bar{X}_{1,n} \leq a \right)^{d_X} da \end{aligned}$$

Applying Bahadur-Rao:

- ① Rate function is $J(a^* - \varepsilon)$

Proof: Upper bound

Integrating by parts we obtain an upper bound

$$\begin{aligned} & e^{-nd_Y l_Y(a^* - \varepsilon)} \mathbb{P} \left(\bar{X}_{1,n} \leq a^* - \varepsilon \right)^{d_X} \\ & + \int_{\mathbb{E} Y}^{a^* - \varepsilon} nd_Y l'_Y(a) e^{-nd_Y l_Y(a)} \mathbb{P} \left(\bar{X}_{1,n} \leq a \right)^{d_X} da \end{aligned}$$

Applying Bahadur-Rao:

- ① Rate function is $J(a^* - \varepsilon)$
- ② $\int_{\mathbb{E} Y}^{a^* - \varepsilon} n^{1-d_X/2} d_Y l'_Y(a) (-C_X(a))^{d_X} e^{-nJ(a)} da$

Proof: Upper bound

Integrating by parts we obtain an upper bound

$$\begin{aligned} & e^{-nd_Y l_Y(a^* - \varepsilon)} \mathbb{P} \left(\bar{X}_{1,n} \leq a^* - \varepsilon \right)^{d_X} \\ & + \int_{\mathbb{E} Y}^{a^* - \varepsilon} n d_Y l'_Y(a) e^{-nd_Y l_Y(a)} \mathbb{P} \left(\bar{X}_{1,n} \leq a \right)^{d_X} da \end{aligned}$$

Applying Bahadur-Rao:

- ① Rate function is $J(a^* - \varepsilon)$

- ② $\underbrace{\int_{\mathbb{E} Y}^{a^* - \varepsilon} n^{1-d_X/2} d_Y l'_Y(a) (-C_X(a))^{d_X} e^{-nJ(a)} da}_{f(a)}$

Proof: Upper bound

Because $J(a) \geq J'(a^* - \varepsilon)(a - a^* + \varepsilon) + J(a^* - \varepsilon)$ for any a ,

$$\begin{aligned} & \int_{\mathbb{E}Y}^{a^* - \varepsilon} n^{1-d_X/2} f(a) e^{-nJ(a)} da \\ & \leq e^{-nJ(a^* - \varepsilon)} n^{1-d_X/2} \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) e^{-nJ'(a^* - \varepsilon)(a - a^* + \varepsilon)} da \end{aligned}$$

Proof: Upper bound

Because $J(a) \geq J'(a^* - \varepsilon)(a - a^* + \varepsilon) + J(a^* - \varepsilon)$ for any a ,

$$\begin{aligned} & \int_{\mathbb{E}Y}^{a^* - \varepsilon} n^{1-d_X/2} f(a) e^{-nJ(a)} da \\ & \leq e^{-nJ(a^* - \varepsilon)} n^{1-d_X/2} \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) e^{-nJ'(a^* - \varepsilon)(a - a^* + \varepsilon)} da \end{aligned}$$

Proof: Upper bound

Because $J(a) \geq J'(a^* - \varepsilon)(a - a^* + \varepsilon) + J(a^* - \varepsilon)$ for any a ,

$$\begin{aligned} & \int_{\mathbb{E}Y}^{a^* - \varepsilon} n^{1-d_X/2} f(a) e^{-nJ(a)} da \\ & \leq e^{-nJ(a^* - \varepsilon)} n^{1-d_X/2} \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) e^{-nJ'(a^* - \varepsilon)(a - a^* + \varepsilon)} da \end{aligned}$$

Proof: Upper bound

Because $J(a) \geq J'(a^* - \varepsilon)(a - a^* + \varepsilon) + J(a^* - \varepsilon)$ for any a ,

$$\begin{aligned} & \int_{\mathbb{E}Y}^{a^* - \varepsilon} n^{1-d_X/2} f(a) e^{-nJ(a)} da \\ & \leq e^{-nJ(a^* - \varepsilon)} n^{1-d_X/2} \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) e^{-nJ'(a^* - \varepsilon)(a - a^* + \varepsilon)} da \end{aligned}$$

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) e^{-nJ'(a^* - \varepsilon)(a - (a^* - \varepsilon))} da \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{E}Y}^{a^* - \varepsilon} f(a) da = 0 \end{aligned}$$



A recipe for mixtures



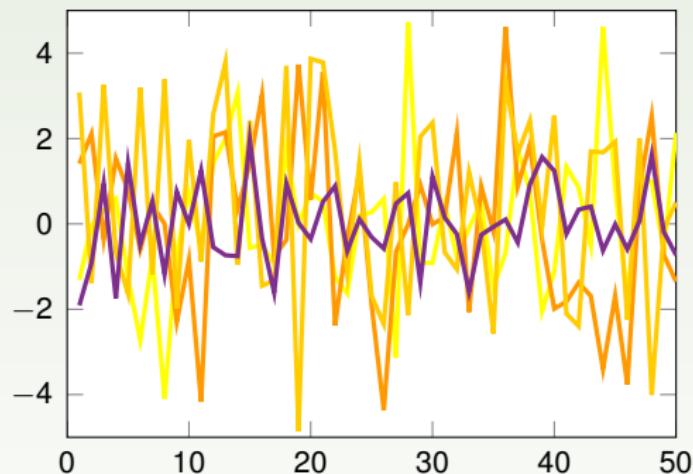
- ① Identify point of concentration a^* ✓
- ② Get exact asymptotics for lower bound $\int_{a^*-\varepsilon}^{a^*+\varepsilon}$, using, e.g.:
 - Bahadur-Rao
 - Taylor approximation
 - ...✓
- ③ Show that first and last summand of upper bound $\int_{-\infty}^{a^*-\varepsilon} + \int_{a^*-\varepsilon}^{a^*+\varepsilon} + \int_{a^*+\varepsilon}^{\infty}$ are negligible (and thus lower and upper bound coincide) ✓

Proposition (K., Mandjes, Taimre)

Suppose that X and Y are continuous, and fulfil the conditions of Bahadur-Rao. Then,

$$\alpha_{d_X, d_Y}(n) \sim n^{(1-d_X-d_Y)/2} K(a^*) \sqrt{\frac{2\pi}{J''(a^*)}} e^{-nJ(a^*)}.$$

Application I: Anomaly identification



Goal: identify the anomalous process (the one with known target distribution G), no prior belief which one it is

Application I: Anomaly identification

We declare process i to be anomalous based on a total of dn observations

$$\mathcal{L}_i(n) := \frac{1}{n} \sum_{k=1}^n \log \frac{g(x_i(k) | x_i(1), \dots, x_i(k-1))}{f(x_i(k) | x_i(1), \dots, x_i(k-1))} > \mathcal{L}_j(n).$$

Application I: Anomaly identification

We declare process i to be anomalous based on a total of dn observations

$$\mathcal{L}_i(n) := \frac{1}{n} \sum_{k=1}^n \log \frac{g(x_i(k) | x_i(1), \dots, x_i(k-1))}{f(x_i(k) | x_i(1), \dots, x_i(k-1))} > \mathcal{L}_j(n).$$

Label anomalous process with 1. The event of a false selection, i.e. of declaring the wrong process to be anomalous, based on dn samples is given by

$$\left\{ \mathcal{L}_1(n) < \max_{i \in \{2, \dots, d\}} \mathcal{L}_i(n) \right\}.$$

Application II: Order statistics

If $i^* := \arg \min_{i \in \{1, \dots, d-k+1\}} J_{d-i+1, k+i-1}(a^*)$, is unique, then we have:

$$\begin{aligned} & \mathbb{P}\left(\exists i \in \{1, \dots, d-k+1\} : \bar{X}_{(i),n} \leq \bar{Y}_{(k+i-1),n}\right) \\ & \sim \binom{d}{k+i^*-1} \binom{d}{d-i^*+1} \alpha_{d-i^*+1, k+i^*-1}(n) \end{aligned}$$

Application II: Order statistics

If $i^* := \arg \min_{i \in \{1, \dots, d-k+1\}} J_{d-i+1, k+i-1}(a^*)$, is unique, then we have:

$$\begin{aligned} & \mathbb{P}\left(\exists i \in \{1, \dots, d-k+1\} : \bar{X}_{(i),n} \leq \bar{Y}_{(k+i-1),n}\right) \\ & \sim \binom{d}{k+i^*-1} \binom{d}{d-i^*+1} \color{red}{\alpha_{d-i^*+1, k+i^*-1}(n)} \end{aligned}$$

The proof relies on the fact that $\mathbb{P}\left(\bar{X}_{(i),n} \leq \bar{Y}_{(i+k-1),n}\right)$ can be written as

$$\binom{d}{k+i-1} \binom{d}{d-i+1} \mathbb{P}\left(\min_{j \in \{1, \dots, k+i-1\}} \bar{Y}_{j,n} > \max_{j \in \{i, \dots, d\}} \bar{X}_{j,n},\right. \\ \left. \min_{j \in \{1, \dots, k+i-1\}} \bar{Y}_{j,n} \geq \max_{j \in \{k+i, \dots, d\}} \bar{Y}_{j,n}, \max_{j \in \{i, \dots, d\}} \bar{X}_{j,n} \leq \min_{j \in \{1, \dots, i-1\}} \bar{X}_{j,n}\right).$$

Mixed Poisson example

$$P_n(a) := \mathbb{P} \left(\text{Pois} \left(n\bar{X}_{n^\alpha} \right) \geq na \right) \sim ?$$

Mixed Poisson example

$$P_n(a) := \mathbb{P} \left(\text{Pois} \left(n\bar{X}_{n^\alpha} \right) \geq na \right) \sim ?$$

Point of concentration:

if $\alpha > 1$: n^α large compared to n , so \bar{X}_{n^α} close to mean,
and Poisson random variable takes value a

Mixed Poisson example

$$P_n(a) := \mathbb{P} \left(\text{Pois} \left(n\bar{X}_{n^\alpha} \right) \geq na \right) \sim ?$$

Point of concentration:

if $\alpha > 1$: n^α large compared to n , so \bar{X}_{n^α} close to mean,
and Poisson random variable takes value a

if $\alpha < 1$: n^α small compared to n , so large deviation caused
by \bar{X}_{n^α} being close to a

Proposition (Heemskerk,K.,Mandjes)

For $\alpha > 3$, as $n \rightarrow \infty$,

$$P_n(a) \sim e^{-n I_{\text{Pois}(\nu)}(a)} \frac{C_{\text{Pois}(\nu)}(a)}{\sqrt{n}}.$$

For $\alpha < \frac{1}{3}$, as $n \rightarrow \infty$,

$$P_n(a) \sim e^{-n^\alpha I_X(a)} \frac{C_X(a)}{n^{\alpha/2}}.$$

Importance sampling

Goal: Estimate $\mathbb{P}(\bar{Y}_n \geq a)$, where $a > \mathbb{E} Y$

Idea: To make sampling more efficient, change the measure such that event of interest is not rare

Importance sampling

Goal: Estimate $\mathbb{P}(\bar{Y}_n \geq a)$, where $a > \mathbb{E} Y$

Idea: To make sampling more efficient, change the measure such that event of interest is not rare

$$\mathbb{P}(\bar{Y}_n \geq a) = \int_{-\infty}^{\infty} \mathbb{1} \left\{ \sum_i y_i \geq na \right\} \underbrace{\frac{p(\mathbf{y})}{q(\mathbf{y})}}_{=:L} q(\mathbf{y}) d\mathbf{y}$$

Importance sampling

Goal: Estimate $\mathbb{P}(\bar{Y}_n \geq a)$, where $a > \mathbb{E} Y$

Idea: To make sampling more efficient, change the measure such that event of interest is not rare

$$\mathbb{P}(\bar{Y}_n \geq a) = \int_{-\infty}^{\infty} \mathbb{1}\left\{\sum_i y_i \geq na\right\} \underbrace{\frac{p(\mathbf{y})}{q(\mathbf{y})}}_{=:L} q(\mathbf{y}) d\mathbf{y}$$

Recall from proof of Bahadur-Rao that the *exponentially tilted* measure $q(y) = \frac{e^{\theta^* y}}{M(y)} p(y)$ seems to be a good idea.

Importance sampling for $\alpha_{d_X, d_Y}(n)$

Exponentially tilted measures:

$$q_X(x) = \frac{e^{\theta_X(a^*)x}}{M_X(\theta_X(a^*))} p_X(x), \quad q_Y(y) = \frac{e^{\theta_Y(a^*)y}}{M_Y(\theta_Y(a^*))} p_Y(y)$$

Is this a good choice?

Asymptotic efficiency

Variance of the estimator:

$$\mathbb{E}_q \left[\left(L \mathbf{1} \left\{ \bar{Y}_n \geq a \right\} \right)^2 \right] - \mathbb{P} \left(\bar{Y}_n \geq a \right)^2 \geq 0$$

Asymptotic efficiency

Variance of the estimator:

$$\mathbb{E}_q \left[\left(L \mathbb{1} \left\{ \bar{Y}_n \geq a \right\} \right)^2 \right] - \mathbb{P} \left(\bar{Y}_n \geq a \right)^2 \geq 0$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_q \left[L^2 \mathbb{1} \left\{ \bar{Y}_n \geq a \right\} \right] \geq -2I(a)$$

Asymptotic efficiency

Variance of the estimator:

$$\mathbb{E}_q \left[\left(L \mathbb{1} \left\{ \bar{Y}_n \geq a \right\} \right)^2 \right] - \mathbb{P} \left(\bar{Y}_n \geq a \right)^2 \geq 0$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_q \left[L^2 \mathbb{1} \left\{ \bar{Y}_n \geq a \right\} \right] \geq -2I(a)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_q \left[L^2 \mathbb{1} \left\{ \bar{Y}_n \geq a \right\} \right] \leq -2I(a)$$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]\end{aligned}$$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]\end{aligned}$$

- $I'(a) = \theta'(a)a + \theta(a) - \frac{M'(\theta(a))}{M(\theta(a))}\theta'(a) = \theta(a)$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]\end{aligned}$$

- $I'(a) = \theta'(a)a + \theta(a) - \frac{M'(\theta(a))}{M(\theta(a))}\theta'(a) = \theta(a)$
- $0 = J'(a^*) = d_X I'_X(a^*) + d_Y I'_Y(a^*) = d_X \theta_X(a^*) + d_Y \theta_Y(a^*)$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]\end{aligned}$$

- $I'(a) = \theta'(a)a + \theta(a) - \frac{M'(\theta(a))}{M(\theta(a))}\theta'(a) = \theta(a)$
- $0 = J'(a^*) = d_X I'_X(a^*) + d_Y I'_Y(a^*) = d_X \theta_X(a^*) + d_Y \theta_Y(a^*)$
- $d_X \theta_X(a^*) = -d_Y \theta_Y(a^*)$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]\end{aligned}$$

$$\begin{aligned}-\theta_Y(a^*) \sum_{i=1}^{d_Y} \sum_{k=1}^n Y_{i,k} &= -d_Y \theta_Y(a^*) \frac{1}{d_Y} \sum_{i=1}^{d_Y} \sum_{k=1}^n Y_{i,k} \\ &\leq -d_Y \theta_Y(a^*) \frac{1}{d_X} \sum_{j=1}^{d_X} \sum_{\ell=1}^n X_{j,\ell} = d_X \theta_X(a^*) \frac{1}{d_X} \sum_{j=1}^{d_X} \sum_{\ell=1}^n X_{j,\ell}\end{aligned}$$

Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \underbrace{\mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]}_{\leq 1}\end{aligned}$$

$$\begin{aligned}\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}} [L^2 \mathbb{1}\{\dots\}] \\ &\leq 2d_X \Lambda_X(\theta_X(a^*)) + 2d_Y \Lambda_Y(\theta_Y(a^*)) \\ &= -2a^* [d_X \theta_X(a^*) + d_Y \theta_Y(a^*)] + 2d_X \Lambda_X(\theta_X(a^*)) + 2d_Y \Lambda_Y(\theta_Y(a^*)) \\ &= -2J(a^*).\end{aligned}$$

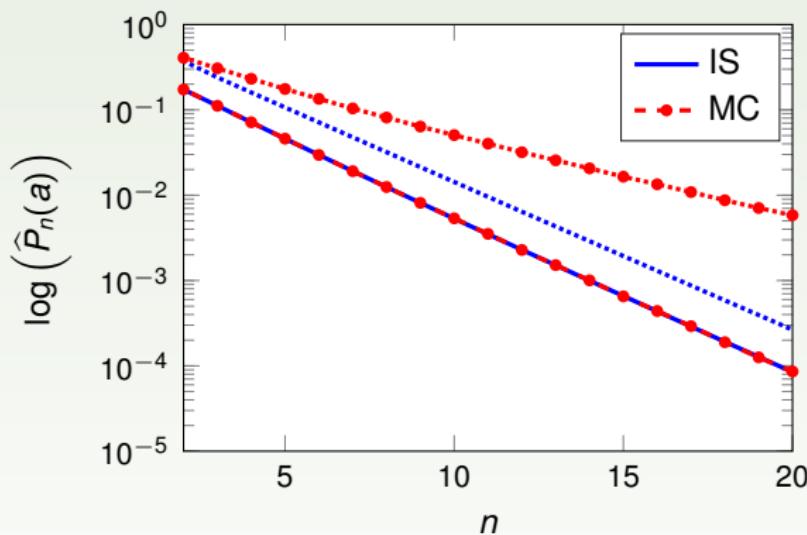
Asymptotic efficiency for $\alpha_{d_X, d_Y}(n)$

$$\begin{aligned}\mathbb{E}_q(L^2 \mathbb{1}\{\dots\}) &= (M_X(\theta(a^*)))^{2nd_X} (M_Y(\theta(a^*)))^{2nd_Y} \\ &\times \underbrace{\mathbb{E}_q \left[e^{-2\theta_X(a^*) \sum_{i=1}^{d_X} \sum_{k=1}^n X_{i,k}} e^{-2\theta_Y(a^*) \sum_{j=1}^{d_Y} \sum_{\ell=1}^n Y_{j,\ell}} \mathbb{1}\{\dots\} \right]}_{\leq 1}\end{aligned}$$

$$\begin{aligned}\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}} [L^2 \mathbb{1}\{\dots\}] \\ &\leq 2d_X \Lambda_X(\theta_X(a^*)) + 2d_Y \Lambda_Y(\theta_Y(a^*)) \\ &= -2a^* [d_X \theta_X(a^*) + d_Y \theta_Y(a^*)] + 2d_X \Lambda_X(\theta_X(a^*)) + 2d_Y \Lambda_Y(\theta_Y(a^*)) \\ &= -2J(a^*).\end{aligned}$$



Example



Logarithmic importance sampling (IS) and crude Monte Carlo (MC) estimators for $P_n(2)$, where $X_i \sim \text{Exp}(1)$, $\alpha = 2$. Upper bounds of sample confidence intervals are indicated by dotted lines (width of intervals is inflated by a factor 10^3).

Conclusions

- Be cautious with logarithmic asymptotics
- To prove exact asymptotics for mixture distributions, we often do not need the change of measure if we know the point of concentration
- Asymptotically efficient importance sampling estimators are suggested by large deviations

Conclusions

- Be cautious with logarithmic asymptotics
- To prove exact asymptotics for mixture distributions, we often do not need the change of measure if we know the point of concentration
- Asymptotically efficient importance sampling estimators are suggested by large deviations

Thank you!