Aggregated Tests Based on Supremal Divergence Estimators for non-Regular Statistical Models

Jean-Patrick Baudry Michel Broniatowski Cyril Thommeret

LPSM – Sorbonne Université – Paris

August 31st 2023

CASM (Csiszar-Ali-Silvey-Morimoto) φ-Divergences

- $\varphi : \mathbb{R} \to \mathbb{R}^+$ differentiable and strictly convex with $\varphi(1) = 0$
- P and P^* probability measures
- φ-divergence between P and P^*:

$$D_\varphi(P, P^*) = \begin{cases} \int \varphi(\frac{dP}{dP^*})\, dP^* \text{ if } P << P^* \\ +\infty \text{ otherwise} \end{cases}$$

- $D_\varphi(P, P^*) = 0 \iff P = P^*$
- examples:
  - for $\varphi(x) = xlogx - x + 1$, D_φ is the Kullback-Leibler divergence
  - for $\varphi(x) = \frac{1}{2}(x-1)^2$, D_φ is the χ^2-divergence

Dual Representation of φ-Divergences

$$D_\varphi(P, P^*) = \begin{cases} \int \varphi(\frac{dP}{dP^*})\, dP^* \text{ if } P << P^* \\ +\infty \text{ otherwise} \end{cases}$$

- $\mathcal{F}$ some class of borelian real valued functions
- For any P such that

$$\begin{cases} \int |f| dP < \infty \text{ for any } f \in \mathcal{F} \\ D_\varphi(P, P^*) < \infty \\ \varphi'(\frac{dP}{dP^*}) \in \mathcal{F} \end{cases}$$

it holds

$$D_\varphi(P, P^*) = \max_{f \in \mathcal{F}} \int f\, dP - \int \varphi^*(f)\, dP^*,$$

where $\varphi^* : t \in \mathbb{R} \mapsto \sup_{x \in \mathbb{R}} tx - \varphi(x)$.
The supremum is uniquely attained at $f = \varphi'(\frac{dP}{dP^*})$.

[Liese and Vajda, 2006, Broniatowski and Keziou, 2006]

Statistical Setting

- $\{f_1(\,.\,;\theta_1) : \theta_1 \in \Theta_1\}$, $\Theta_1 \subset \mathbb{R}^p$, and $\{f_2(\,.\,;\theta_2) : \theta_2 \in \Theta_2\}$, $\Theta_2 \subset \mathbb{R}^q$ probability density families with respect to a σ-finite measure λ on $(\mathcal{X}, \mathcal{B})$
- for any $(\pi, \theta) \in \Theta$ with $\theta = (\theta_1, \theta_2)$,

$$g_{\pi,\theta} = (1-\pi) f_1(\,.\,;\theta_1) + \pi f_2(\,.\,;\theta_2)$$

- $X_1, \ldots, X_n$ i.i.d. sample with distribution $P^* := g_{\pi^*,\theta^*}.\lambda$ (unknown parameters $(\pi^*, \theta^*) \in \Theta$)
- Aim: inference on π^*

Supremal Estimator of the Parameters in a Mixture Model

$$g_{\pi,\theta} = (1-\pi)f_1(\,.\,;\theta_1) + \pi f_2(\,.\,;\theta_2)$$

- g a probability density (escort parameter) such that

$$\forall(\pi,\theta) \in \Theta, \begin{cases} Supp(g) \subset Supp(g_{\pi,\theta}) \\ \int \left|\varphi'(\frac{g}{g_{\pi,\theta}})\right| g\, d\lambda < \infty \end{cases}$$

- For any $(\pi,\theta) \in \Theta$, define

$$m_{\pi,\theta} : x \in \mathcal{X} \mapsto \int \varphi'\left(\frac{g}{g_{\pi,\theta}}\right) g\, d\lambda - \varphi^* \circ \varphi'\left(\frac{g}{g_{\pi,\theta}}\right)(x)$$

- Dual representation of the divergence:

$$D_\varphi(g.\lambda, g_{\pi^*,\theta_1^*,\theta_2^*}.\lambda) = \max_{(\pi,\theta_1)\in]a,b[\times\Theta_1} \mathbb{E}_{P^*}[m_{\pi,\theta_1,\theta_2^*}(X)]$$

$$(\pi^*,\theta_1^*) = \operatorname*{argmax}_{(\pi,\theta_1)\in]a,b[\times\Theta_1} \mathbb{E}_{P^*}[m_{\pi,\theta_1,\theta_2^*}(X)]$$

Supremal Estimator of the Parameters in a Mixture Model

$$g_{\pi,\theta} = (1-\pi) f_1(\,.\,;\theta_1) + \pi f_2(\,.\,;\theta_2)$$

- g a probability density (escort parameter)
- $(\pi^*, \theta_1^*) = \mathrm{argmax}_{(\pi,\theta_1)\in]a,b[\times\Theta_1} \mathbb{E}_{P^*}[m_{\pi,\theta_1,\theta_2^*}(X)]$
- θ_2 is not estimated
- Supremal estimator of (π^*, θ_1^*):

$$\forall\theta_2 \in \Theta_2, (\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)\in]a,b[\times\Theta_1}{\mathrm{argmax}} \frac{1}{n}\sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i)$$

- $P^* \to \mathbb{P}_n$ is a legitimate substitution when $\theta_2 = \theta_2^*$ or $\pi^* = 0$ (since $X_1, \ldots, X_n \overset{i.i.d.}{\sim} P^* = g_{\pi^*,\theta^*}.\lambda$)

Supremal Estimator of the Parameters in a Mixture Model

$$g_{\pi,\theta} = (1-\pi) f_1(\,.\,;\theta_1) + \pi f_2(\,.\,;\theta_2)$$

- g a probability density (escort parameter)
- $(\pi^*, \theta_1^*) = \text{argmax}_{(\pi,\theta_1)\in]a,b[\times\Theta_1} \mathbb{E}_{P^*}[m_{\pi,\theta_1,\theta_2^*}(X)]$
- θ_2 is not estimated
- Supremal estimator of (π^*, θ_1^*):

$$\forall \theta_2 \in \Theta_2, (\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)\in]a,b[\times\Theta_1}{\text{argmax}} \frac{1}{n}\sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i)$$

- $P^* \to \mathbb{P}_n$ is a legitimate substitution when $\theta_2 = \theta_2^*$ or $\pi^* = 0$ (since $X_1, \ldots, X_n \overset{i.i.d.}{\sim} P^* = g_{\pi^*,\theta^*}.\lambda$)

Supremal Estimator of the Parameters in a Mixture Model

$$g_{\pi,\theta} = (1-\pi) f_1(\,.\,;\theta_1) + \pi f_2(\,.\,;\theta_2)$$

- g a probability density (escort parameter)
- $(\pi^*, \theta_1^*) = \text{argmax}_{(\pi,\theta_1)\in]a,b[\times\Theta_1} \mathbb{E}_{P^*}[m_{\pi,\theta_1,\theta_2^*}(X)]$
- θ_2 is not estimated
- Supremal estimator of (π^*, θ_1^*):

$$\forall \theta_2 \in \Theta_2, (\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)\in]a,b[\times\Theta_1}{\text{argmax}} \frac{1}{n}\sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i)$$

- $P^* \to \mathbb{P}_n$ is a legitimate substitution when $\theta_2 = \theta_2^*$ or $\pi^* = 0$ (since $X_1, \ldots, X_n \overset{i.i.d.}{\sim} P^* = g_{\pi^*,\theta^*}.\lambda$)

Consistency and Asymptotic Distribution of the Supremal Estimator

$$(\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i) \quad \text{with} \quad \pi \in]a, b[\ni 0$$

If $\pi^* = 0$, for any $\theta_2, \theta_2' \in \Theta_2$, and under regularity conditions,

$$\begin{cases} \hat{\pi}(\theta_2) \xrightarrow{a.s.} 0 \\ \hat{\theta}_1(\theta_2) \xrightarrow{a.s.} \theta_1^* \end{cases}$$

and

$$\begin{pmatrix} \sqrt{\frac{n}{a_n}}(\hat{\pi}(\theta_2) - \pi^*) \\ \sqrt{\frac{n}{a_n'}}(\hat{\pi}(\theta_2') - \pi^*) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} 1 & \frac{b}{\sqrt{aa'}} \\ \frac{b}{\sqrt{aa'}} & 1 \end{pmatrix}\right)$$

where a_n (resp. a_n') depends only on θ_2 (resp. θ_2') and the sample but a, a', and b depend on the (unknown) distribution P^*.

Consistency and Asymptotic Distribution of the Supremal Estimator

$$(\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i) \quad \text{with} \quad \pi \in]a, b[\ni 0$$

If $\pi^* = 0$, for any $\theta_2, \theta_2' \in \Theta_2$, and under regularity conditions,

$$\begin{cases} \hat{\pi}(\theta_2) \xrightarrow{a.s.} 0 \\ \hat{\theta}_1(\theta_2) \xrightarrow{a.s.} \theta_1^* \end{cases}$$

and

$$\begin{pmatrix} \sqrt{\frac{n}{a_n}}(\hat{\pi}(\theta_2) - \pi^*) \\ \sqrt{\frac{n}{a_n'}}(\hat{\pi}(\theta_2') - \pi^*) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} 1 & \frac{b}{\sqrt{aa'}} \\ \frac{b}{\sqrt{aa'}} & 1 \end{pmatrix}\right)$$

where a_n (resp. a_n') depends only on θ_2 (resp. θ_2') and the sample but a, a', and b depend on the (unknown) distribution P^*.

Consistency and Asymptotic Distribution of the Supremal Estimator

$$(\hat{\pi}(\theta_2), \hat{\theta}_1(\theta_2)) \in \underset{(\pi,\theta_1)}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} m_{\pi,\theta_1,\theta_2}(X_i) \quad \text{with} \quad \pi \in]a, b[\ni 0$$

If $\pi^* = 0$, for any $\theta_2, \theta_2' \in \Theta_2$, and under regularity conditions,

$$\begin{cases} \hat{\pi}(\theta_2) \xrightarrow{a.s.} 0 \\ \hat{\theta}_1(\theta_2) \xrightarrow{a.s.} \theta_1^* \end{cases}$$

and

$$\begin{pmatrix} \sqrt{\frac{n}{a_n}}(\hat{\pi}(\theta_2) - \pi^*) \\ \sqrt{\frac{n}{a_n'}}(\hat{\pi}(\theta_2') - \pi^*) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} 1 & \frac{b}{\sqrt{aa'}} \\ \frac{b}{\sqrt{aa'}} & 1 \end{pmatrix}\right)$$

where a_n (resp. a_n') depends only on θ_2 (resp. θ_2') and the sample but a, a', and b depend on the (unknown) distribution P^*.

Test the Number of Components in a Mixture Model

Aim: based on a realisation of the sample, test the hypothesis
$H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$
Test statistic:

$$T_n = \sup_{\theta_2} \sqrt{\frac{n}{a_n}} \hat{\pi}(\theta_2)$$

Reject H_0 if T_n large $\iff$ if there is enough evidence of a second component for some θ_2 if $T_n > t_\alpha$ such that for $\pi^* = 0$ and any $\theta_1^* \in \Theta_1$,

$$P^*(T_n > t_\alpha) \leq \alpha.$$

- with G centred Gaussian process s.t. $\mathrm{Cov}(G_{\theta_2}, G_{\theta_2'}) = \frac{b}{\sqrt{aa'}}$,

$$T_n \xrightarrow{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- with G^n centred Gaussian process s.t. $\mathrm{Cov}(G^n_{\theta_2}, G^n_{\theta_2'}) = \frac{b_n}{\sqrt{a_n a_n'}}$ and Θ_2^δ finite δ-grid,

$$\sup_{\theta_2 \in \Theta_2^\delta} G^n_{\theta_2} \xrightarrow[\substack{n \to \infty \\ \delta \to 0}]{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

Test the Number of Components in a Mixture Model

Aim: based on a realisation of the sample, test the hypothesis
$H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$
Test statistic:

$$T_n = \sup_{\theta_2} \sqrt{\frac{n}{a_n}} \hat{\pi}(\theta_2)$$

Reject H_0 if T_n large $\iff$ if there is enough evidence of a second component for some θ_2 if $T_n > t_\alpha$ such that for $\pi^* = 0$ and any $\theta_1^* \in \Theta_1$,

$$P^*(T_n > t_\alpha) \leq \alpha.$$

- with G centred Gaussian process s.t. $\mathrm{Cov}(G_{\theta_2}, G_{\theta_2'}) = \frac{b}{\sqrt{aa'}}$,

$$T_n \xrightarrow{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- with G^n centred Gaussian process s.t. $\mathrm{Cov}(G^n_{\theta_2}, G^n_{\theta_2'}) = \frac{b_n}{\sqrt{a_n a_n'}}$ and Θ_2^δ finite δ-grid,

$$\sup_{\theta_2 \in \Theta_2^\delta} G^n_{\theta_2} \xrightarrow[\substack{n\to\infty \\ \delta\to 0}]{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

Test the Number of Components in a Mixture Model

Aim: based on a realisation of the sample, test the hypothesis
$H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$
Test statistic:

$$T_n = \sup_{\theta_2} \sqrt{\frac{n}{a_n}} \hat{\pi}(\theta_2)$$

Reject H_0 if T_n large $\iff$ if there is enough evidence of a second component for some θ_2 if $T_n > t_\alpha$ such that for $\pi^* = 0$ and any $\theta_1^* \in \Theta_1$,

$$P^*(T_n > t_\alpha) \leq \alpha.$$

- with G centred Gaussian process s.t. $\mathrm{Cov}(G_{\theta_2}, G_{\theta_2'}) = \frac{b}{\sqrt{aa'}}$,

$$T_n \xrightarrow{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- with G^n centred Gaussian process s.t. $\mathrm{Cov}(G^n_{\theta_2}, G^n_{\theta_2'}) = \frac{b_n}{\sqrt{a_n a_n'}}$ and Θ_2^δ finite δ-grid,

$$\sup_{\theta_2 \in \Theta_2^\delta} G^n_{\theta_2} \xrightarrow[\substack{n \to \infty \\ \delta \to 0}]{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

Test the Number of Components in a Mixture Model

Aim: based on a realisation of the sample, test the hypothesis $H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$

Test statistic:

$$T_n = \sup_{\theta_2} \sqrt{\frac{n}{a_n}} \hat{\pi}(\theta_2)$$

Reject H_0 if T_n "large": if $T_n > t_\alpha$ such that for $\pi^* = 0$ and any $\theta_1^* \in \Theta_1$,

$$P^*(T_n > t_\alpha) \leq \alpha.$$

- with G centred Gaussian process s.t. $\mathrm{Cov}(G_{\theta_2}, G_{\theta_2'}) = \frac{b}{\sqrt{aa'}}$,

$$T_n \xrightarrow{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- with G^n centred Gaussian process s.t. $\mathrm{Cov}(G^n_{\theta_2}, G^n_{\theta_2'}) = \frac{b_n}{\sqrt{a_n a_n'}}$ and Θ_2^δ finite δ-grid,

$$\sup_{\theta_2 \in \Theta_2^\delta} G^n_{\theta_2} \xrightarrow[\substack{n \to \infty \\ \delta \to 0}]{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- G depends on P^* but G^n depends on the sample only.

Test the Number of Components in a Mixture Model

Aim: based on a realisation of the sample, test the hypothesis
$H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$
Test statistic:

$$T_n = \sup_{\theta_2} \sqrt{\frac{n}{a_n}} \hat{\pi}(\theta_2)$$

Reject H_0 if T_n "large": if $T_n > t_\alpha$ such that for $\pi^* = 0$ and any $\theta_1^* \in \Theta_1$,

$$P^*(T_n > t_\alpha) \leq \alpha.$$

- with G centred Gaussian process s.t. $\mathrm{Cov}(G_{\theta_2}, G_{\theta_2'}) = \frac{b}{\sqrt{aa'}}$,

$$T_n \xrightarrow{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- with G^n centred Gaussian process s.t. $\mathrm{Cov}(G^n_{\theta_2}, G^n_{\theta_2'}) = \frac{b_n}{\sqrt{a_n a_n'}}$ and Θ_2^δ finite δ-grid,

$$\sup_{\theta_2 \in \Theta_2^\delta} G^n_{\theta_2} \xrightarrow[\substack{n \to \infty \\ \delta \to 0}]{\mathcal{L}} \sup_{\theta_2 \in \Theta_2} G_{\theta_2}$$

- G depends on P^* but G^n depends on the sample only.

Test the Number of Components in a Mixture Model

Algorithm

Input : φ, g, $\{f_1(\, . \, ; \theta_1) : \theta_1 \in \Theta_1\}$, $\{f_2(\, . \, ; \theta_2) : \theta_2 \in \Theta_2\}$, K, Θ_2^δ, $p \in [0, 1]$, $(x_1, \dots, x_n)$

1. let $t = \sup_{\theta_2 \in \Theta_2} \sqrt{\frac{n}{a_n(\theta_2)}} \hat{\pi}(\theta_2)$
2. for $k \in \{1, \dots, K\}$
   2.1 sample $(G_t)_{t \in \Theta_2^\delta} \sim \mathcal{N}\big(0, (\frac{b_n(t,t')}{\sqrt{a_n(t)a_n(t')}})_{t,t' \in \Theta_2^\delta}\big)$
   2.2 let $\tilde{t}_k = \max_{t \in \Theta_2^\delta} g_t$
3. if $t \geq$ empirical_quantile$((\tilde{t}_k)_{k \in \{1,\dots,K\}}, 1 - p)$ reject H_0 else don't reject H_0 vs H_1

Example: Mixtures of Uniform Distributions

Let F be a (known) continuous cdf.
Test
H_0: the observations can be modelled as a sample from F
versus
H_1: a proportion π^* (unknown) of this data has been obtained by discarding all values larger than some $c^* \in \mathbb{R}$ (unknown)

- If $Y \sim F$ and $\eta^* = F(c^*)$, the cdf of $Y|Y \leq c^*$ is $G : t \in [0,1] \mapsto \frac{1}{\eta^*} F(t) \mathbb{1}_{0 \leq F(t) \leq \eta^*} + \mathbb{1}_{\eta^* < F(t)}$.
- If $X \sim (1-\pi^*)F + \pi^* G$ with $\pi^* \in [0,1]$, then $F(X) \sim (1-\pi^*)\mathcal{U}([0,1]) + \pi^* \mathcal{U}([0,\eta^*])$.
- Test H_0: $\pi^* = 0$ vs H_1: $\pi^* > 0$ for the observations $F(x_1), \ldots, F(x_n)$ in the model

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]}.$$

Example: Mixtures of Uniform Distributions

Let F be a (known) continuous cdf.
Test
H_0: the observations can be modelled as a sample from F
versus
H_1: a proportion π^* (unknown) of this data has been obtained by discarding all values larger than some $c^* \in \mathbb{R}$ (unknown)

- If $Y \sim F$ and $\eta^* = F(c^*)$, the cdf of $Y|Y \leq c^*$ is $G : t \in [0,1] \mapsto \frac{1}{\eta^*} F(t) \mathbb{1}_{0 \leq F(t) \leq \eta^*} + \mathbb{1}_{\eta^* < F(t)}$.
- If $X \sim (1-\pi^*)F + \pi^* G$ with $\pi^* \in [0,1]$, then $F(X) \sim (1-\pi^*)\mathcal{U}([0,1]) + \pi^* \mathcal{U}([0,\eta^*])$.
- Test H_0: $\pi^* = 0$ vs H_1: $\pi^* > 0$ for the observations $F(x_1), \ldots, F(x_n)$ in the model

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]}.$$

Example: Mixtures of Uniform Distributions

- $f_1 = \mathbb{1}_{[0,1]}$, $f_2(.;\eta) = \frac{1}{\eta}\mathbb{1}_{[0,\eta]}$ ($\theta_2 = \eta$, no θ_1 to estimate)
- for any η,
$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]}$$
- consider the modified Kullback-Leibler divergence: $\phi : x \in \mathbb{R} \mapsto -\log x + x - 1$
- The associated supremal estimator is:
$$\hat{\pi}(\eta) = \begin{cases} p_- + \frac{\eta}{\eta-1}p_+ = \frac{p_- - \eta}{1-\eta} \text{ if } n_- > 0 \\ 0 \text{ if } n_- = 0 \end{cases}$$
with $p_- = \frac{n_-}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq \eta}$ and $p_+ = 1 - p_-$
- $\hat{\pi}(\eta)$ is the usual maximum likelihood estimator when $\hat{\pi}(\eta) \geq 0 (\Leftrightarrow p_- \geq \eta)$

Example: Mixtures of Uniform Distributions

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]} \qquad p_- = \frac{n_-}{n} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{X_i \leq \eta} \qquad p_+ = 1 - p_-$$

- $X_1, \dots, X_n \sim g_{\pi^*,\eta^*}$. Test $H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$.
- $H_1 = \cup_\eta H_1(\eta)$ where $H_1(\eta) : X_1, \dots, X_n \sim g_{\pi^*,\eta}$ with $\pi^* > 0$
- usual likelihood ratio test:

$$\begin{aligned} LRTS &= 2\log \frac{\sup_\eta \prod_{i=1}^n g_{\hat{\pi}(\eta)\vee 0,\eta}(X_i)}{\prod_{i=1}^n \mathbb{1}_{[0,1]}(X_i)} \\ &= 2\log \sup_{\eta \leq p_-} \left(\frac{p_-}{\eta}\right)^{n_-}\left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1 \end{aligned}$$

Non-regular model: under H_0, $LRTS \xrightarrow{\mathcal{L}} \sup_\eta \xi_\eta^2 \mathbb{1}_{\xi_\eta > 0}$

Based on Feng & McCulloch (1992), we may consider to extend the range of π:

$$\begin{aligned} LRTSe &= 2\log \frac{\sup_\eta \prod_{i=1}^n g_{\hat{\pi}(\eta),\eta}(X_i)}{\prod_{i=1}^n \mathbb{1}_{[0,1]}(X_i)} \\ &= 2\log \sup_{\eta} \left(\frac{p_-}{\eta}\right)^{n_-}\left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1 \end{aligned}$$

Example: Mixtures of Uniform Distributions

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]} \qquad p_- = \frac{n_-}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \le \eta} \qquad p_+ = 1 - p_-$$

- $X_1, \ldots, X_n \sim g_{\pi^*,\eta^*}$. Test $H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$.
- $H_1 = \cup_\eta H_1(\eta)$ where $H_1(\eta) : X_1, \ldots, X_n \sim g_{\pi^*,\eta}$ with $\pi^* > 0$
- usual likelihood ratio test:

$$LRTS = 2\log \frac{\sup_\eta \prod_{i=1}^n g_{\hat{\pi}(\eta)\vee 0,\eta}(X_i)}{\prod_{i=1}^n \mathbb{1}_{[0,1]}(X_i)} = 2\log \sup_{\eta \le p_-} \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1$$

Non-regular model: under H_0, $LRTS \xrightarrow{\mathcal{L}} \sup_\eta \xi_\eta^2 \mathbb{1}_{\xi_\eta > 0}$
Based on Feng & McCulloch (1992), we may consider to extend the range of π:

$$LRTSe = 2\log \frac{\sup_\eta \prod_{i=1}^n g_{\hat{\pi}(\eta),\eta}(X_i)}{\prod_{i=1}^n \mathbb{1}_{[0,1]}(X_i)} = 2\log \sup_{\eta} \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1$$

Example: Mixtures of Uniform Distributions

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]} \qquad p_- = \frac{n_-}{n} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{X_i \le \eta} \qquad p_+ = 1 - p_-$$

- $X_1, \ldots, X_n \sim g_{\pi^*,\eta^*}$. Test $H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$.
- $H_1 = \cup_\eta H_1(\eta)$ where $H_1(\eta) : X_1, \ldots, X_n \sim g_{\pi^*,\eta}$ with $\pi^* > 0$
- usual likelihood ratio test:

$$LRTS = \sup_{\eta \le p_-} \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1 \xrightarrow{\mathcal{L}} \sup_\eta \ \xi_\eta^2 \mathbb{1}_{\xi_\eta > 0}$$

$$LRTSe = \sup_\eta \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1$$

- our test is not a likelihood ratio test:

$$T = \sup_\eta \sqrt{\frac{n}{a_n}}\,\hat{\pi}(\eta) = \sup_\eta \sqrt{n}\,\frac{p_- - \eta}{\sqrt{p_- p_+}} \xrightarrow{\mathcal{L}} \sup_\eta \xi_\eta$$

Example: Mixtures of Uniform Distributions

$$g_{\pi,\eta} = (1-\pi)\mathbb{1}_{[0,1]} + \frac{\pi}{\eta}\mathbb{1}_{[0,\eta]} \qquad p_- = \frac{n_-}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \le \eta} \qquad p_+ = 1 - p_-$$

- $X_1, \ldots, X_n \sim g_{\pi^*,\eta^*}$. Test $H_0 : \pi^* = 0$ vs $H_1 : \pi^* > 0$.
- $H_1 = \cup_\eta H_1(\eta)$ where $H_1(\eta) : X_1, \ldots, X_n \sim g_{\pi^*,\eta}$ with $\pi^* > 0$
- usual likelihood ratio test:

$$LRTS = \sup_{\eta \le p_-} \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1 \xrightarrow{\mathcal{L}} \sup_\eta \ \xi_\eta^2 \mathbb{1}_{\xi_\eta > 0}$$

$$LRTSe = \sup_\eta \left(\frac{p_-}{\eta}\right)^{n_-} \left(\frac{p_+}{1-\eta}\right)^{n_+} \vee 1$$

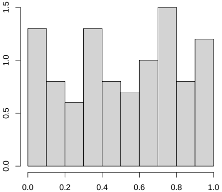
- our test is not a likelihood ratio test:

$$T = \sup_\eta \sqrt{\frac{n}{a_n}}\,\hat{\pi}(\eta) = \sup_\eta \sqrt{n}\,\frac{p_- - \eta}{\sqrt{p_- p_+}} \xrightarrow{\mathcal{L}} \sup_\eta \xi_\eta$$

Mixtures of Uniforms: Numerical Experiments

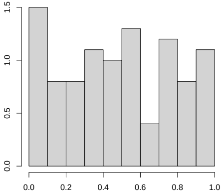
$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} g_{\pi^*,\eta^*} = (1-\pi^*)\mathbb{1}_{[0,1]} + \frac{\pi^*}{\eta^*}\mathbb{1}_{[0,\eta^*]}$$

$$n = 100$$



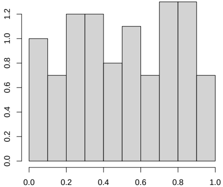
$\pi^* = 0$

Under H_0



$\pi^* = 0.4$
$\eta^* = 0.8$

Under H_1



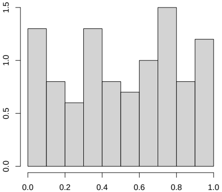
$\pi^* = 0.1$
$\eta^* = 0.8$

Under H_1

Mixtures of Uniforms: Numerical Experiments

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} g_{\pi^*,\eta^*} = (1-\pi^*)\mathbb{1}_{[0,1]} + \frac{\pi^*}{\eta^*}\mathbb{1}_{[0,\eta^*]}$$

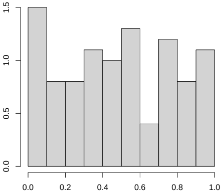
$$n = 100$$



$\pi^* = 0$

Under H_0

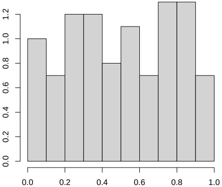
$(\mathbb{P}(X \leq 0.8) = 0.8)$



$\pi^* = 0.4$
$\eta^* = 0.8$

Under H_1

$(\mathbb{P}(X \leq 0.8) = 0.88)$



$\pi^* = 0.1$
$\eta^* = 0.8$

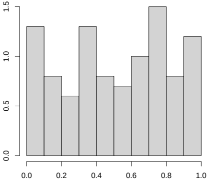
Under H_1

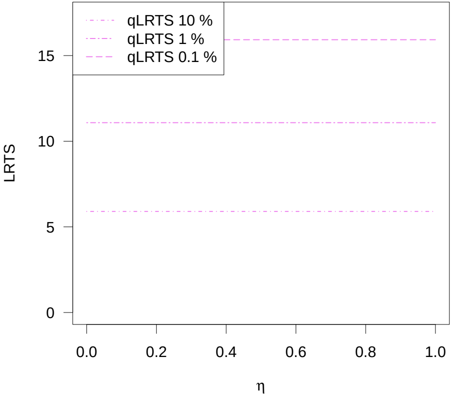
$(\mathbb{P}(X \leq 0.8) = 0.82)$

Mixtures of Uniforms: Numerical Experiments

Under H_0

$$\pi^* = 0$$
$$n = 100$$





q_{LRTS} by Monte-Carlo such that

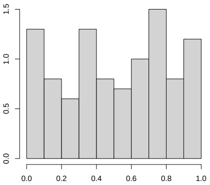
$$P_{\pi^*=0}(LRTS > q_{\text{LRTS}}) \le p$$

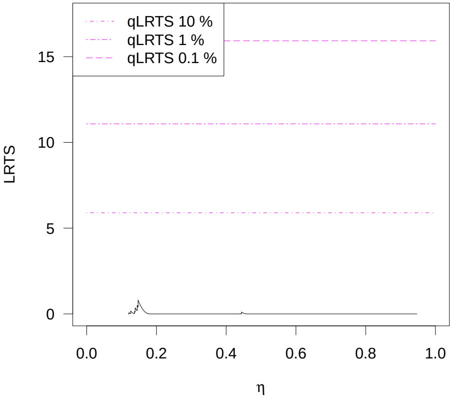
$(p = 0.1, 0.01, 0.001)$

Mixtures of Uniforms: Numerical Experiments

Under H_0

$$\pi^* = 0$$
$$n = 100$$





q_{LRTS} by Monte-Carlo such that

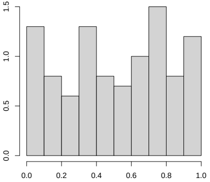
$$P_{\pi^*=0}(LRTS > q_{\text{LRTS}}) \leq p$$

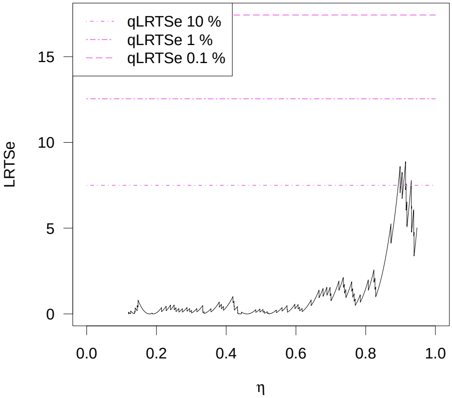
$$(p = 0.1, 0.01, 0.001)$$

Mixtures of Uniforms: Numerical Experiments

Under H_0

$$\pi^* = 0$$
$$n = 100$$





q_{LRTSe} by Monte-Carlo such that

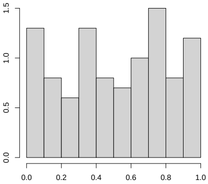
$$P_{\pi^*=0}(LRTSe > q_{\text{LRTSe}}) \leq p$$

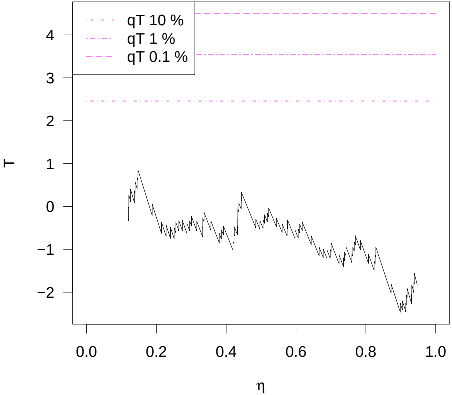
$$(p = 0.1, 0.01, 0.001)$$

Mixtures of Uniforms: Numerical Experiments

Under H_0

$$\pi^* = 0$$
$$n = 100$$





q_T by Monte-Carlo such that

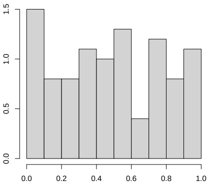
$$P_{\pi^*=0}(T > q_T) \leq p$$

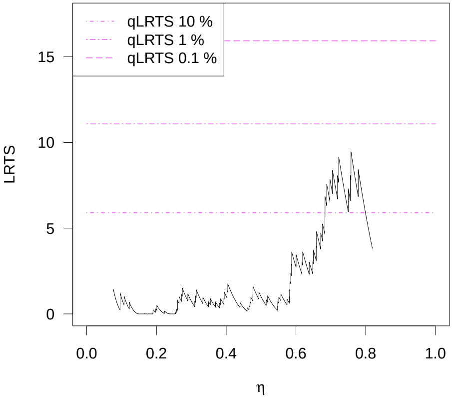
$(p = 0.1, 0.01, 0.001)$

Mixtures of Uniforms: Numerical Experiments

Under H_1

$$\pi^* = 0.4$$
$$\eta^* = 0.8$$
$$n = 100$$





q_{LRTS} by Monte-Carlo such that

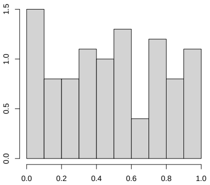
$$P_{\pi^*=0}(LRTS > q_{\text{LRTS}}) \leq p$$

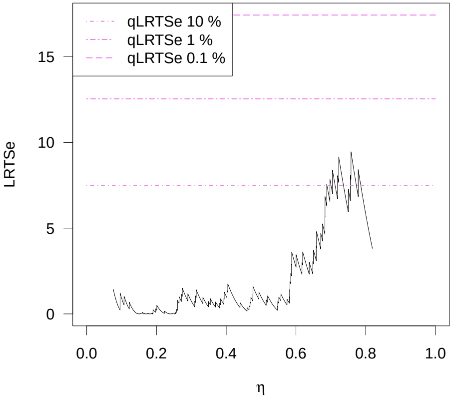
$$(p = 0.1, 0.01, 0.001)$$

Mixtures of Uniforms: Numerical Experiments

Under H_1

$$\pi^* = 0.4$$
$$\eta^* = 0.8$$
$$n = 100$$





q_{LRTSe} by Monte-Carlo such that

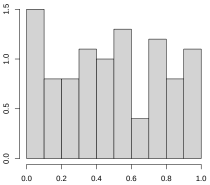
$$P_{\pi^*=0}(LRTSe > q_{\mathrm{LRTSe}}) \leq p$$

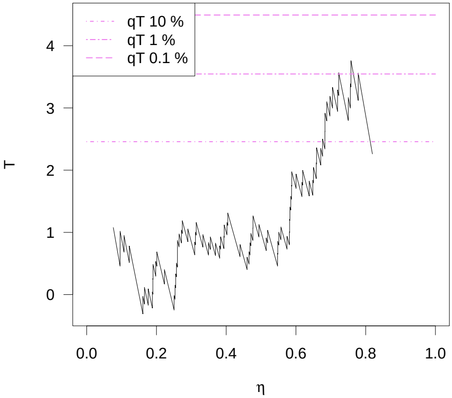
$$(p = 0.1, 0.01, 0.001)$$

Mixtures of Uniforms: Numerical Experiments

Under H_1

$$\pi^* = 0.4$$
$$\eta^* = 0.8$$
$$n = 100$$





q_T by Monte-Carlo such that

$$P_{\pi^*=0}(T > q_T) \leq p$$

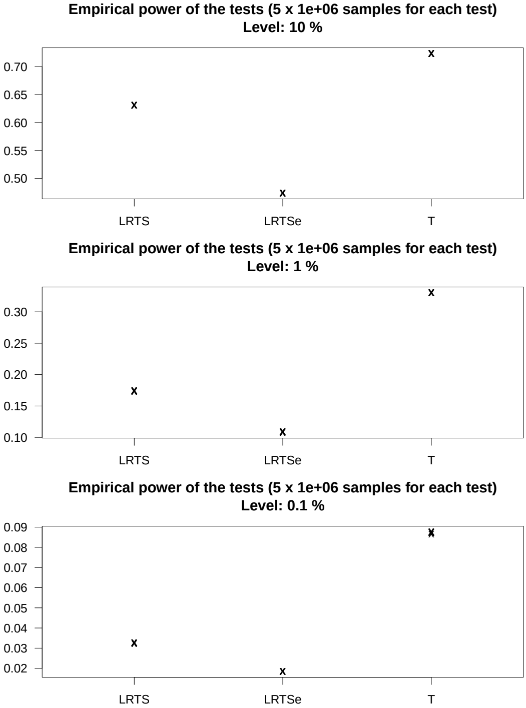
$(p = 0.1, 0.01, 0.001)$

Monte-Carlo approximation of the probability to reject H_0 under H_1 for each test

$\pi^* = 0.4$

$\eta^* = 0.8$

$n = 100$

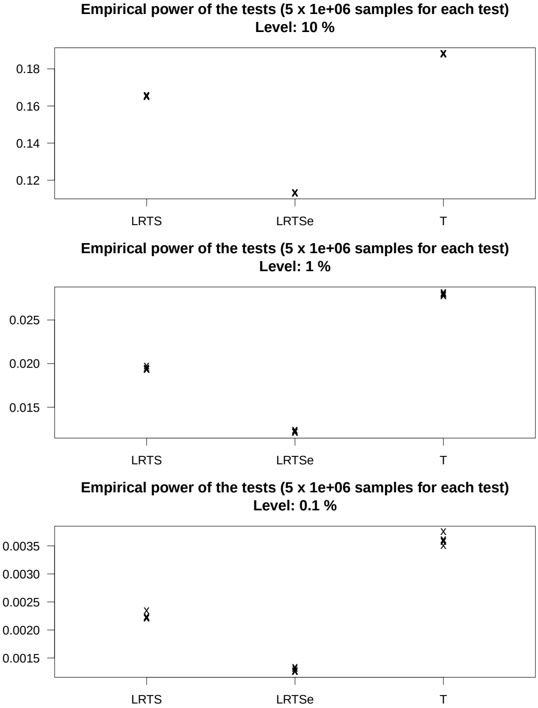


Monte-Carlo approximation of the probability to reject H_0 under H_1 for each test

$\pi^* = 0.1$

$\eta^* = 0.8$

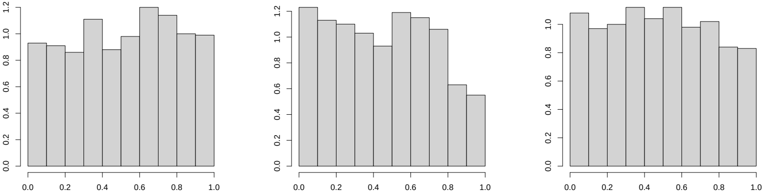
$n = 100$



Mixtures of Uniforms: Numerical Experiments

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} g_{\pi^*,\eta^*} = (1-\pi^*)\mathbb{1}_{[0,1]} + \frac{\pi^*}{\eta^*}\mathbb{1}_{[0,\eta^*]}$$

$$n = 1000$$



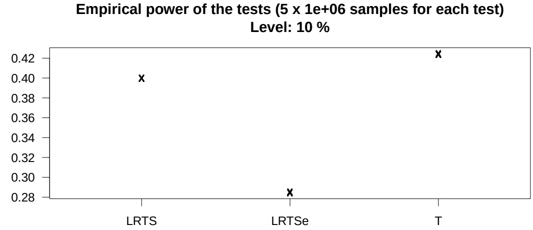
$\pi^* = 0$	$\pi^* = 0.4$ $\eta^* = 0.8$	$\pi^* = 0.1$ $\eta^* = 0.8$
Under H_0	Under H_1	Under H_1

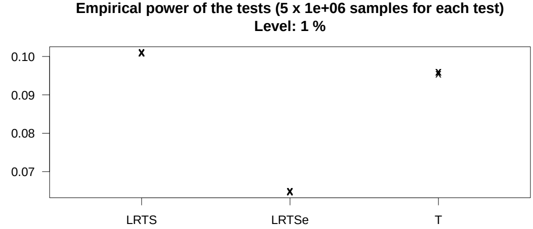
Monte-Carlo approximation of the probability to reject H_0 under H_1 for each test

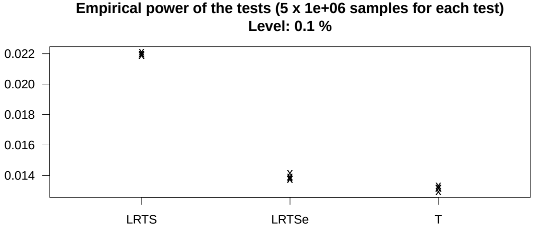
$\pi^* = 0.1$

$\eta^* = 0.8$

$n = 1000$





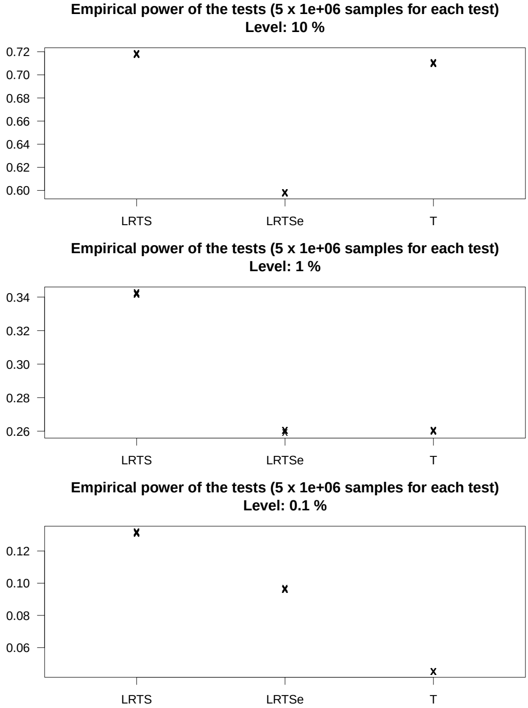


Monte-Carlo approximation of the probability to reject H_0 under H_1 for each test

$\pi^* = 0.1$

$\eta^* = 0.6$

$n = 1000$



Conclusion and Perspectives

- Work in progress
- Other divergences φ, choice of g: robustness?
- Situations with unknown θ_1 can be addressed
- Mixtures components in neighbourhoods of given families of probability measures

Thank you for your attention!

Broniatowski, M. and Keziou, A. (2006).
Minimization of φ-divergences on sets of signed measures.
Studia Scientiarum Mathematicarum Hungarica, 43(4):403–442.

Liese, F. and Vajda, I. (2006).
On divergences and informations in statistics and information theory.
IEEE Transactions on Information Theory, 52(10):4394–4412.