f-Divergences and Applications to Privacy

Mario Diaz (IIMAS – UNAM)

Based on joint work with Shahab Asoodeh (McMaster) & Flavio Calmon (Harvard)

Slides available at *https://www.mariodiaztorres.com/*

Outline

Outline

Outline

f-Divergences and Contraction

Probability Simplex

Probability Simplex

Def. Given n 2 N, the probability simplex in R^n is the set defined as

$$P([n]) \coloneqq fp\ 2\ \mathrm{R}^n : p_1;:::;p_n \quad 0;p_1 + \quad + p_n = 1g:$$

Probability Simplex

Def. Given $n \ 2$ N, the probability simplex in R^n is the set defined as

$$P([n]) \coloneqq fp \ 2 \ \mathrm{R}^n : p_1;:::;p_n \quad 0;p_1 + \quad + p_n = 1g:$$

Probability Simplex

Def. Given $n \ 2$ N, the probability simplex in R^n is the set defined as

$$P([n]) \coloneqq fp \ 2 \ \mathrm{R}^n : p_1;:::;p_n \quad 0;p_1 + \quad + p_n = 1g:$$

$P([n])$ encodes the set of probability distributions over $f1;:::;ng$.

Probability Simplex

Def. Given $n \ 2$ N, the probability simplex in R^n is the set defined as

$$P([n]) \coloneqq fp \ 2 \ \mathrm{R}^n : p_1;:::;p_n \quad 0;p_1 + \quad + p_n = 1g:$$

$P([n])$ encodes the set of probability distributions over $f1;:::;ng$.

From a statistical perspective, the geometry of $P([n])$ is not obvious...

f-Divergences I: Definition

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \to \mathbb{R}$ be a convex function with $f(1) = 0$.

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \; ! \; \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in X} f\left(\frac{P(x)}{Q(x)}\right) Q(x):$$

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \to \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in X} f\left(\frac{P(x)}{Q(x)}\right) Q(x).$$

Examples of f-divergences

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \; ! \; \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in X} f\left(\frac{P(x)}{Q(x)}\right) Q(x):$$

Examples of f-divergences

Total Variation Distance: $f(t) = \frac{1}{2}|t - 1|$

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \; ! \; \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in X} f\left(\frac{P(x)}{Q(x)}\right) Q(x):$$

Examples of f-divergences

Total Variation Distance: $f(t) = \frac{1}{2}|t - 1|$

Hellinger Divergence: $f(t) = \frac{t - 1}{1}$

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \; ! \; \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in \mathcal{X}} f\left(\frac{P(x)}{Q(x)}\right) Q(x):$$

Examples of f-divergences

Total Variation Distance: $f(t) = \frac{1}{2}|t - 1|$

Hellinger Divergence: $f(t) = \dfrac{t - 1}{1}$

KL-Divergence: $f(t) = t \log(t)$

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences I: Definition

Notation. We let $f : (0; 1) \; ! \; \mathbb{R}$ be a convex function with $f(1) = 0$.

Def. Given pmfs P and Q, the f-divergence of P w.r.t. Q is defined as

$$D_f(P\|Q) = \sum_{x \in X} f\left(\frac{P(x)}{Q(x)}\right) Q(x):$$

Examples of f-divergences

Total Variation Distance: $f(t) = \frac{1}{2}|t - 1|$

Hellinger Divergence: $f(t) = \frac{(t-1)^2}{t}$

KL-Divergence: $f(t) = t \log(t)$

Some f-divergences have useful operational interpretations!

I. Csiszár. "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, 1967.

f-Divergences II: Basic Properties

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(P\|Q) \quad 0$

f-Divergences II: Basic Properties

$D_f(P\|Q) \quad 0$

If f is strictly convex at 1, then

$D_f(P\|Q) = 0 \;\;,\;\; P = Q.$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(P\|Q) \quad 0$

If f is strictly convex at 1, then

$D_f(P\|Q) = 0 \;\;,\;\; P = Q.$

$(P;Q)$?! $D_f(P\|Q)$ is convex.

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(PkQ) \quad 0$

If f is strictly convex at 1, then

$D_f(PkQ) = 0 \;\;,\;\; P = Q.$

$(P;Q) \;\overline{?}\; D_f(PkQ)$ is convex.

Notation. We denote Markov kernels or privacy mechanisms by

$$K(yjx) = \mathrm{P}(Y = yjX = x):$$

$$X \;\; ! \boxed{K} ! \;\; Y$$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(PkQ) \quad 0$

If f is strictly convex at 1, then

$D_f(PkQ) = 0 \ , \ P = Q.$

$(P;Q)$ 7! $D_f(PkQ)$ is convex.

Notation. We denote Markov kernels or privacy mechanisms by

$$K(yjx) = \mathrm{P}(Y = yjX = x):$$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(PkQ) \quad 0$

If f is strictly convex at 1, then

$D_f(PkQ) = 0 \;\;,\;\; P = Q.$

$(P;Q)$ 7! $D_f(PkQ)$ is convex.

Notation. We denote Markov kernels or privacy mechanisms by

$$K(yjx) = \mathrm{P}(Y = yjX = x):$$

Given a pmf P, we let KP be defined as

$$KP(y) = \sum_{x2X} P(x)K(yjx):$$

$$P \;\; ! \;\; \boxed{K} \;\; ! \;\; KP$$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(PkQ) \quad 0$

If f is strictly convex at 1, then

$D_f(PkQ) = 0 \;\;,\;\; P = Q.$

$(P;Q)$ 7! $D_f(PkQ)$ is convex.

Notation. We denote Markov kernels or privacy mechanisms by

$$K(yjx) = \mathrm{P}(Y = yjX = x):$$

Given a pmf P, we let KP be defined as

$$KP(y) = \sum_{x2X} P(x)K(yjx):$$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences II: Basic Properties

$D_f(PkQ) \quad 0$

If f is strictly convex at 1, then

$D_f(PkQ) = 0 \;\;,\;\; P = Q.$

$(P;Q)$ 7! $D_f(PkQ)$ is convex.

Notation. We denote Markov kernels or privacy mechanisms by

$$K(yjx) = \mathrm{P}(Y = yjX = x):$$

Given a pmf P, we let KP be defined as

$$KP(y) = \sum_{x2X} P(x)K(yjx):$$

$P \;!\; \boxed{K} \;!\; KP$

$Q \;!\; \boxed{K} \;!\; KQ$

Data Processing Inequality (DPI): $D_f(KPkKQ) \quad D_f(PkQ)$

Y. Polyanskiy & Y. Wu. *Lecture Notes on Information Theory*, 2014.

f-Divergences III: Hockey-Stick Divergence

f-Divergences III: Hockey-Stick Divergence

Def. For $\quad 2\ (0;\ 1\ )$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad\ t)_+ & < 1; \\ (t\quad\ )_+ & > 1; \end{cases}$$

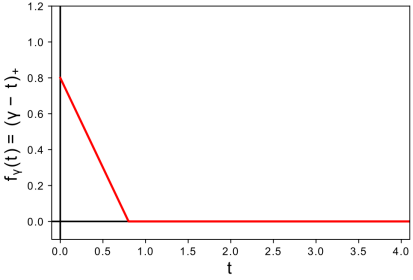
where $(x)_+ = \max\{0; x\}$.

f-Divergences III: Hockey-Stick Divergence

Def. For $2\ (0;\ 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.

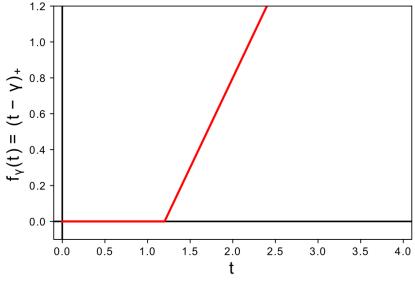


f-Divergences III: Hockey-Stick Divergence

Def. For $2\ (0;\ 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.

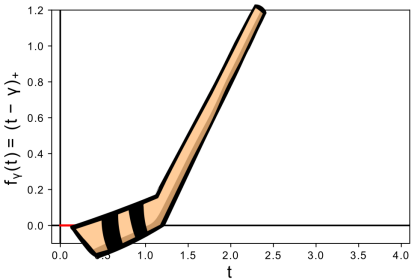


f-Divergences III: Hockey-Stick Divergence

Def. For 2 $(0; 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f(t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.

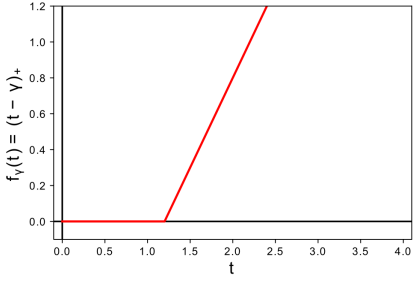


f-Divergences III: Hockey-Stick Divergence

Def. For $2\ (0;\ 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.



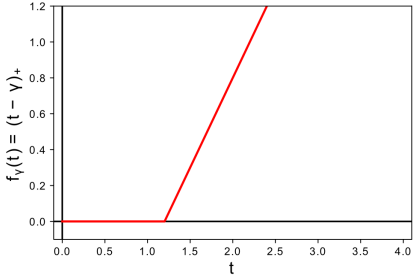
f-Divergences III: Hockey-Stick Divergence

Def. For $2\ (0;\ 1\ )$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.

$$\lim_{\ 1} E\ (P\|Q) = \mathrm{TV}(P;Q)$$



f-Divergences III: Hockey-Stick Divergence

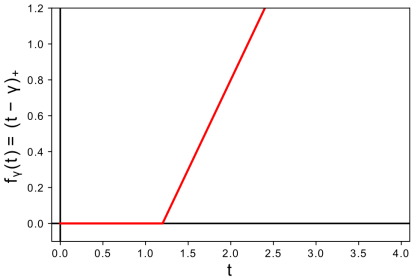
Def. For $2\ (0;\ 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.

$$\lim_{\ 1} E\ (P\|Q) = \mathrm{TV}(P;Q)$$

$$< 1\text{: } E\ (P\|Q) = \sup_{A\ \ X}[\ \ Q(A) \quad P(A)]$$

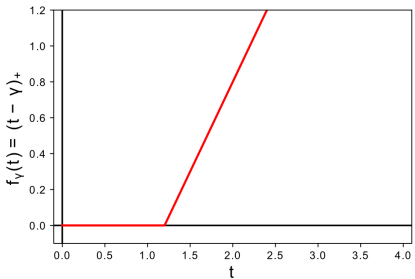


f-Divergences III: Hockey-Stick Divergence

Def. For $2\ (0;\ 1)$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.



$$\lim_{\ 1} E\ (P\|Q) = \mathrm{TV}(P; Q)$$

$$< 1\text{: } E\ (P\|Q) = \sup_{A\ \ X} [\ \ Q(A) \quad P(A)]$$

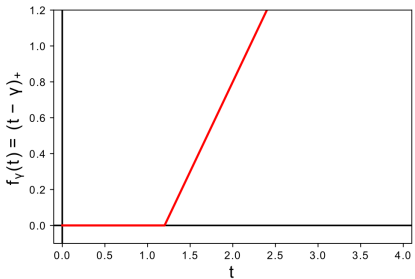
$$> 1\text{: } E\ (P\|Q) = \sup_{A\ \ X} [P(A) \quad Q(A)]$$

f-Divergences III: Hockey-Stick Divergence

Def. For $\quad 2\ (0;\ 1\ )$, the Hockey-Stick divergence is the f-divergence defined by

$$f\ (t) = \begin{cases} (\quad t)_+ & < 1; \\ (t \quad)_+ & > 1; \end{cases}$$

where $(x)_+ = \max\{0; x\}$.



$$\lim_{\ \to 1} E\ (P\|Q) = \mathrm{TV}(P;Q)$$

$$< 1\text{: } E\ (P\|Q) = \sup_{A\ \ X}[\ \ Q(A) \quad P(A)]$$

$$> 1\text{: } E\ (P\|Q) = \sup_{A\ \ X}[P(A) \quad Q(A)]$$

If f is twice differentiable, then

$$D_f(P\|Q) = \int_0^{\ 1} E\ (P\|Q) f''(\ )\mathrm{d}\ .$$

f-Divergences IV: Relation between f-Divergences

f-Divergences IV: Relation between f-Divergences

Pinsker's Inequality: $\mathrm{TV}(P;Q) \quad \sqrt{\frac{D_{\mathrm{KL}}(P\|Q)}{2}}$

f-Divergences IV: Relation between f-Divergences

Pinsker's Inequality: $\mathrm{TV}(P;Q) \quad \sqrt{\frac{D_{\mathrm{KL}}(P\|Q)}{2}}$

If $H^2(P;Q)$ is the squared Hellinger distance (i.e., $f(t) = (1 - \sqrt{t})^2$), then

$$\frac{1}{2}H^2(P;Q) \quad \mathrm{TV}(P;Q) \quad H(P;Q):$$

f-Divergences IV: Relation between f-Divergences

Pinsker's Inequality: $\mathrm{TV}(P;Q) \quad \sqrt{\frac{D_{\mathrm{KL}}(P\|Q)}{2}}$

If $H^2(P;Q)$ is the squared Hellinger distance (i.e., $f(t) = (1 - \sqrt{t})^2$), then

$$\frac{1}{2}H^2(P;Q) \quad \mathrm{TV}(P;Q) \quad H(P;Q):$$

There are many relations between f-divergences...

f-Divergences IV: Relation between f-Divergences

Pinsker's Inequality: $\mathrm{TV}(P;Q) \quad \sqrt{\frac{D_{\mathrm{KL}}(P\|Q)}{2}}$

If $H^2(P;Q)$ is the squared Hellinger distance (i.e., $f(t) = (1 - \sqrt{t})^2$), then

$$\frac{1}{2}H^2(P;Q) \quad \mathrm{TV}(P;Q) \quad H(P;Q):$$

There are many relations between f-divergences...

...but there is a systematic form to obtain them!

f-Divergences V: Joint Range Strategy

P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q); D_g(P\|Q)) : P;Q \text{ pmfs on some alphabet } X\};$$

$$R_2(f;g) := \{(D_f(P\|Q); D_g(P\|Q)) : P;Q \text{ pmfs on } \{0;1\}\}:$$

P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on some alphabet } X\};$$
$$R_2(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on } \{0,1\}\}.$$

Thm. $R(f;g) = \mathrm{co}(R_2(f;g))$

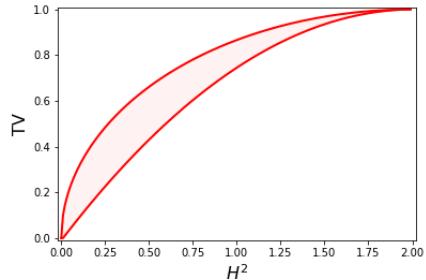
P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on some alphabet } X\};$$
$$R_2(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on } \{0,1\}\}.$$

Thm. $R(f;g) = \text{co}(R_2(f;g))$



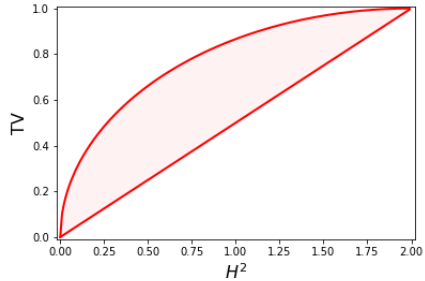
P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on some alphabet } X\};$$
$$R_2(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on } \{0,1\}\}.$$

Thm. $R(f;g) = \text{co}(R_2(f;g))$



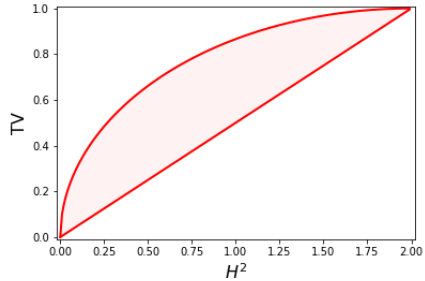
P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on some alphabet } X\};$$
$$R_2(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on } \{0,1\}\}.$$

Thm. $R(f;g) = \text{co}(R_2(f;g))$



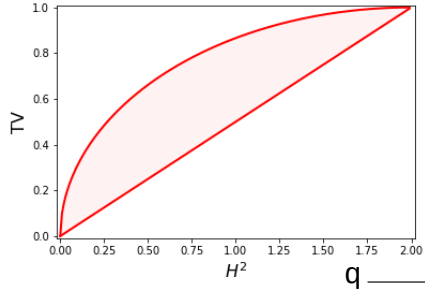
P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences V: Joint Range Strategy

Notation. Given two f-divergences D_f and D_g, their joint range is defined as

$$R(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on some alphabet } X\};$$
$$R_2(f;g) := \{(D_f(P\|Q), D_g(P\|Q)) : P, Q \text{ pmfs on } \{0,1\}\}.$$

Thm. $R(f;g) = \mathrm{co}(R_2(f;g))$



$\frac{1}{2}H^2(P;Q)$ $\quad$ $\mathrm{TV}(P;Q)$ $\quad$ $H(P;Q)$ $\quad$ 1 $\quad$ $H^2(P;Q)=4$

P. Harremoës & I. Vajda. "On pairs of f-divergences and their joint range," *IEEE Transactions on Information Theory*, 2011.

f-Divergences VI: Strong Data Processing Inequality

R. Ahlswede & P. Gács. "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, 1976.
R.L. Dobrushin. "Central limit theorem for nonstationary Markov chains. I," *Theory of Probability & Its Applications*, 1956.

f-Divergences VI: Strong Data Processing Inequality

Def. The contraction coefficient of a Markov kernel K under f-divergence is defined as

$$_f(K) = \sup_{\substack{P;Q: \\ 0<D_f(PkQ)<1}} \frac{D_f(KPkKQ)}{D_f(PkQ)}:$$

R. Ahlswede & P. Gács. "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, 1976.
R.L. Dobrushin. "Central limit theorem for nonstationary Markov chains. I," *Theory of Probability & Its Applications*, 1956.

f-Divergences VI: Strong Data Processing Inequality

Def. The contraction coefficient of a Markov kernel K under f-divergence is defined as

$$_f(K) = \sup_{\substack{P;Q: \\ 0<D_f(PkQ)<1}} \frac{D_f(KPkKQ)}{D_f(PkQ)}:$$

For all $P;Q$, we have $D_f(KPkKQ)$ $_f(K)D_f(PkQ)$.

$$P \;!\; \boxed{K} \;!\; KP$$

$$Q \;!\; \boxed{K} \;!\; KQ$$

R. Ahlswede & P. Gács. "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, 1976.
R.L. Dobrushin. "Central limit theorem for nonstationary Markov chains. I," *Theory of Probability & Its Applications*, 1956.

f-Divergences VI: Strong Data Processing Inequality

Def. The contraction coefficient of a Markov kernel K under f-divergence is defined as

$$_f(K) = \sup_{\substack{P;Q: \\ 0<D_f(PkQ)<1}} \frac{D_f(KPkKQ)}{D_f(PkQ)}:$$

For all $P;Q$, we have $D_f(KPkKQ)$ $_f(K)D_f(PkQ)$.

$$P \;!\; \boxed{K} \;!\; KP$$

$$Q \;!\; \boxed{K} \;!\; KQ$$

Thm. If K is a Markov kernel, then $_{TV}(K) = \sup_{x_1;x_2 2X} TV(K(\,jx_1)kK(\,jx_2))$.

R. Ahlswede & P. Gács. "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, 1976.
R.L. Dobrushin. "Central limit theorem for nonstationary Markov chains. I," *Theory of Probability & Its Applications*, 1956.

f-Divergences VII: Generalized Dobrushin's Formula

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

f-Divergences VII: Generalized Dobrushin's Formula

Notation. For ease of notation, we let $\coloneqq E$.

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

f-Divergences VII: Generalized Dobrushin's Formula

Notation. For ease of notation, we let $\coloneqq {}_{E}$.

Thm. Let $2\ (0; 1]$. If K is a Markov kernel, then

$$(K) = \sup_{x_1;x_2 2X} E\ (K(\,jx_1)kK(\,jx_2)):$$

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

f-Divergences VII: Generalized Dobrushin's Formula

Notation. For ease of notation, we let $\coloneqq {}_{E}$.

Thm. Let $2\ (0; 1\ ]$. If K is a Markov kernel, then

$$(K) = \sup_{x_1;x_2 2X} E\ (K(\ jx_1)kK(\ jx_2)):$$

Inequalities between contraction coefficients

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

f-Divergences VII: Generalized Dobrushin's Formula

Notation. For ease of notation, we let $\coloneqq$ $_E$.

Thm. Let 2 $(0; 1)$. If K is a Markov kernel, then

$$(K) = \sup_{x_1;x_2 2X} E\ (K(\ jx_1)kK(\ jx_2)):$$

Inequalities between contraction coefficients

$$(K) \quad _{\mathrm{TV}}(K) \quad 1 \quad \frac{1 \quad (K)}{} \text{ for all } \quad 1$$

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

f-Divergences VII: Generalized Dobrushin's Formula

Notation. For ease of notation, we let $:= _E$.

Thm. Let $2 (0; 1)$. If K is a Markov kernel, then

$$(K) = \sup_{x_1;x_2 2X} E \ (K(jx_1)kK(jx_2)):$$

Inequalities between contraction coefficients

$$(K) \quad _{TV}(K) \quad 1 \quad \frac{1 \quad (K)}{} \text{ for all } \quad 1$$

$$_f(K) \quad _{TV}(K) \text{ for all } f\text{-divergence}$$

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.
J. Cohen, J.H. Kempermann & G. Zbaganu. *Comparisons of Stochastic Matrices with Applications*, 1998.

Local Di erential Privacy

Big Data, Big Scandals

Big Data, Big Scandals

Big Data, Big Scandals

Big Data, Big Scandals

Big Data, Big Scandals

Big Data, Big Scandals

Big Data, Big Scandals

Ignoring privacy could be problematic...

Big Data, Big Scandals

Ignoring privacy could be problematic...

Different settings, different precise meanings...

Privacy in Surveys

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Privacy in Surveys

Personal Information

Collected Information

Respondent

Curator

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Privacy in Surveys

Personal Information

Collected Information

Respondent

Curator [untrusted]

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Privacy in Surveys

Personal Information

Collected Information

K

Respondent

Curator [untrusted]

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Privacy in Surveys

Personal Information

Collected Information

K

Respondent

Curator [untrusted]

Randomized Response Mechanism

1. Toss a coin.
2. If heads, then answer honestly.
3. If tails, then toss and answer "Yes" if heads, "No" if tails.

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Privacy in Surveys

Personal Information

Collected Information

K

Respondent

Curator [untrusted]

Randomized Response Mechanism

1. Toss a coin.
2. If heads, then answer honestly.
3. If tails, then toss and answer "Yes" if heads, "No" if tails.

Def. For $\quad 0$, the randomized response mechanism is defined by

$K(\,j0) = \mathrm{Ber}(1 \quad p\,) \;\&\; K(\,j1) = \mathrm{Ber}(p\,);$

where $p \;= e\; =(1 + e\;)$.

S.L. Warner. "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, 1965.

Local Di erential Privacy

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Local Di erential Privacy

Notation. Recall that $K(A|x) = P(Y \in A | X = x)$ where $X \to \boxed{K} \to Y$.

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Local Di erential Privacy

Notation. Recall that $K(A|x) = P(Y \ 2\ A|X = x)$ where X ! K ! Y.

Def. Let $\quad 0$ and $\ 2\ [0;1]$. A kernel K is $(\ ;\ )$-LDP if, for every $x_1; x_2\ 2\ X$ and $A\ \ Y$,

$$K(A|x_1) \quad e\ K(A|x_2) + \ :$$

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Local Di erential Privacy

Notation. Recall that $K(A|x) = P(Y \ 2 \ A|X = x)$ where $X \ ! \ \boxed{K} \ ! \ Y$.

Def. Let $\quad 0$ and $\ 2 \ [0; 1]$. A kernel K is $(\ ; \ )$-LDP if, for every $x_1; x_2 \ 2 \ X$ and $A \ \ Y$,

$$K(A|x_1) \quad e \ K(A|x_2) + \ :$$

For small $\ $ and $\ 2 \ [0; 1]$, $jK(y|x_1) \quad K(y|x_2)j \ . \quad + \ .$

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Local Di erential Privacy

Notation. Recall that $K(A|x) = P(Y \ 2\ A|X = x)$ where X ! K ! Y.

Def. Let $\quad 0$ and $\ 2\ [0;1]$. A kernel K is $(\ ;\ )$-LDP if, for every $x_1; x_2\ 2\ X$ and $A\ \ Y$,

$$K(A|x_1) \quad e\ K(A|x_2) + \ :$$

For small $\ $ and $\ 2\ [0;1]$, $jK(y|x_1) \quad K(y|x_2)j$. $\quad +\ \ .$

The likelihood of y given x does not depend much on x.

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Local Di erential Privacy

Notation. Recall that $K(A|x) = P(Y \in A | X = x)$ where $X \to \boxed{K} \to Y$.

Def. Let $\quad 0$ and $\in [0;1]$. A kernel K is $(\ ;\ )$-LDP if, for every $x_1; x_2 \in X$ and $A \quad Y$,

$$K(A|x_1) \quad e \ K(A|x_2) + \ :$$

For small $\quad$ and $\in [0;1]$, $jK(y|x_1) \quad K(y|x_2)j \ . \quad + \ .$

The likelihood of y given x does not depend much on x.

Example. The randomized response mechanism is $(0;\ )$-LDP.

A. Ev mievski, J. Gehrke & R. Srikant. \Limiting privacy breaches in privacy preserving data mining," ACM PODS, 2003.

Three Archetypal Problems in Privacy Analysis

Three Archetypal Problems in Privacy Analysis

Understand privacy from a `fundamental' perspective

Three Archetypal Problems in Privacy Analysis

Understand privacy from a `fundamental' perspective

Assess the cost of privacy in speci c statistical applications

Three Archetypal Problems in Privacy Analysis

Understand privacy from a `fundamental' perspective

Assess the cost of privacy in speci c statistical applications

Compute/estimate the privacy parameters of speci c mechanisms

LDP as Contraction of HS-Divergence

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof.

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Aj x_1) \quad e\ K(Aj x_2)]$

HS-Divergence Representation $E\ (PkQ) = \sup_{A\ X} [P(A) \quad Q(A)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

HS-Divergence Representation $E\ (PkQ) = \sup_{A\ X} [P(A) \quad Q(A)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

, $\sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$

HS-Divergence Representation $E\ (PkQ) = \sup_{A\ X} [P(A) \quad Q(A)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

, $\sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$

HS-Divergence Representation $E\ (PkQ) = \sup_{A\ X} [P(A) \quad Q(A)]$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP ,

$$\sup_{x_1;x_2 \in X} \sup_{A \ Y} [K(A|x_1) \quad e\, K(A|x_2)]$$

,

$$\sup_{x_1;x_2 \in X} E_e\,(K(\ |x_1)\|K(\ |x_2))$$

HS-Divergence Representation: $E\ (P\|Q) = \sup_{A \ X} [P(A) \quad Q(A)]$

Generalized Dobrushin's Formula: $(K) = \sup_{x_1;x_2 \in X} E\ (K(\ |x_1)\|K(\ |x_2))$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

, $\sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$

HS-Divergence Representation: $E\ (PkQ) = \sup_{A\ X} [P(A) \quad Q(A)]$

Generalized Dobrushin's Formula: $(K) = \sup_{x_1;x_2 2X} E\ (K(\ jx_1)kK(\ jx_2))$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $\ _e(K)\quad$.

Proof. K is $(\ ;\ )$-LDP $\quad,\quad \sup_{x_1;x_2 2X}\ \sup_{A\ Y}[K(Ajx_1)\quad e\ K(Ajx_2)]$

$$, \quad \sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$$

$$, \quad \ _e(K)$$

HS-Divergence Representation: $E\ (PkQ) = \sup_{A\ X}[P(A)\quad Q(A)]$

Generalized Dobrushin's Formula: $(K) = \sup_{x_1;x_2 2X} E\ (K(\ jx_1)kK(\ jx_2))$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $\ _e(K)\quad$.

Proof. K is $(\ ;\ )$-LDP $\quad,\quad \sup_{x_1;x_2 2X} \sup_{A\ Y}[K(Ajx_1) \quad e\ K(Ajx_2)]$

$$, \quad \sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$$

$$, \quad \ _e(K)$$

□

HS-Divergence Representation: $E\ (PkQ) = \sup_{A\ X}[P(A) \quad Q(A)]$

Generalized Dobrushin's Formula: $(K) = \sup_{x_1;x_2 2X} E\ (K(\ jx_1)kK(\ jx_2))$

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP , $\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$

, $\sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$

, $_{e}(K)$ □

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP ,

$$\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$$

$$, \quad \sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$$

$$, \quad {}_{e}(K)$$

□

Duchi et al. '13: If K is $(\ ;0)$-LDP, then $D_{KL}(KPkKQ) \quad 2(e \quad 1)^2 kP \quad Qk_{TV}$.

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP ,

$$\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$$

$$, \quad \sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$$

$$, \quad {}_{e}(K)$$

□

Duchi et al. '13: If K is $(\ ;0)$-LDP, then $D_{KL}(KPkKQ) \quad 2(e \quad 1)^2 kP \quad Qk_{TV}$.

\Our main technique... shows that applying di erentially private sampling schemes essentially acts as a contraction on distributions."

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

LDP as Contraction of HS-Divergence

Thm. A privacy mechanism K is $(\ ;\ )$-LDP if and only if $_{e}(K)$.

Proof. K is $(\ ;\ )$-LDP ,

$$\sup_{x_1;x_2 2X} \sup_{A\ Y} [K(Ajx_1) \quad e\ K(Ajx_2)]$$

$$, \quad \sup_{x_1;x_2 2X} E_e\ (K(\ jx_1)kK(\ jx_2))$$

$$, \quad {}_{e}(K)$$

□

Duchi et al. '13: If K is $(\ ;0)$-LDP, then $D_{KL}(KPkKQ) \quad 2(e \quad 1)^2 kP \quad Qk_{TV}$.

\Our main technique... shows that applying di erentially private sampling schemes essentially acts as a contraction on distributions."

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Three Archetypal Problems in Privacy Analysis

Understand privacy from a `fundamental' perspective

Assess the cost of privacy in speci c statistical applications

Compute/estimate the privacy parameters of speci c mechanisms

Binary Hypothesis Testing

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Alternative hypothesis H_1: $X \sim Q$.

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis $H_0 : X \sim P$,

Alternative hypothesis $H_1 : X \sim Q$.

Test. Randomized function $T : \mathcal{X} \to \{0, 1\}$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis $H_0 : X \sim P$,

Alternative hypothesis $H_1 : X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_0)$$

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Alternative hypothesis H_1: $X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_0)$$

False Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_1)$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Alternative hypothesis H_1: $X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_0)$$

False Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_1)$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Alternative hypothesis H_1: $X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_0)$$

False Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_1)$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis H_0: $X \sim P$,

Alternative hypothesis H_1: $X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_0)$$

False Negative Rate

$$(\ ) := P(\ (X) = 0 \mid H_1)$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

Binary Hypothesis Testing

Goal. From a sample of X, choose between

Null hypothesis $H_0 : X \sim P$,

Alternative hypothesis $H_1 : X \sim Q$.

Test. Randomized function $: \mathcal{X} \to \{0, 1\}$

True Negative Rate

$$( ) := P((X) = 0 \mid H_0)$$

False Negative Rate

$$( ) := P((X) = 0 \mid H_1)$$

$$(P; Q) := \inf_{: \ ( )} ( )$$

Cost of LDP in Binary Hypothesis Testing

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let P^n be the distribution of

$$X_1, \ldots, X_n \overset{iid}{\sim} P.$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let $P^{\otimes n}$ be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Stein's Lemma. If $\varepsilon \in (0, 1)$, then

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta_\varepsilon(P^{\otimes n}, Q^{\otimes n})} = D_{\mathrm{KL}}(P \| Q).$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let $P^{\otimes n}$ be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Under (ε, δ)-LDP: $P \to KP$, $Q \to KQ$ &

$$D_{KL}(KP \| KQ) \quad \varepsilon_{KL}(K) D_{KL}(P \| Q).$$

Stein's Lemma. If $\alpha \in (0, 1)$, then

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(P^{\otimes n}, Q^{\otimes n})} = D_{KL}(P \| Q).$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E_γ-divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let $P^{\otimes n}$ be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Stein's Lemma. If $\alpha \in (0,1)$, then

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(P^{\otimes n}, Q^{\otimes n})} = D_{KL}(P \| Q).$$

Under (ε, δ)-LDP: $P \to KP$, $Q \to KQ$ &

$$D_{KL}(KP \| KQ) \leq \eta_{KL}(K) D_{KL}(P \| Q).$$

By the relation between f and β,

$$\eta_{KL}(K) \leq 1 - \frac{1 - \eta_{e^\varepsilon}(K)}{e^\varepsilon}.$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E$_\gamma$-divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let P^n be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Stein's Lemma. If $\alpha \in (0,1)$, then

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta(P^n, Q^n)} = D_{KL}(P\|Q).$$

Under (ε, δ)-LDP: $P \to KP$, $Q \to KQ$ &

$$D_{KL}(KP\|KQ) \le \eta_{KL}(K) D_{KL}(P\|Q).$$

By the relation between f and β,

$$\eta_{KL}(K) \le 1 - \frac{1 - \beta_e(K)}{e^\varepsilon}.$$

Since K is (ε, δ)-LDP, $\beta_e(K) \ge \delta$,

$$\eta_{KL}(K) \le 1 - \frac{1-\delta}{e^\varepsilon}.$$

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E$_\gamma$-divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let $P^{\otimes n}$ be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Stein's Lemma. If $\alpha \in (0,1)$, then

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta(P^{\otimes n}, Q^{\otimes n})} = D_{KL}(P\|Q).$$

Under (ε, δ)-LDP: $P \to KP$, $Q \to KQ$ &

$D_{KL}(KP\|KQ) \quad {}_{KL}(K) D_{KL}(P\|Q)$.

By the relation between $_f$ and ,

$$_{KL}(K) \quad 1 \quad \frac{1 \quad {}_e(K)}{e}.$$

Since K is (ε, δ)-LDP , ${}_e(K) \quad$,

$$_{KL}(K) \quad 1 \quad \frac{1}{e}.$$

Thm. If K is (ε, δ)-LDP, then $\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta(KP^{\otimes n}, KQ^{\otimes n})} \quad 1 \quad \frac{1}{e} \quad D_{KL}(P\|Q)$.

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.

S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Cost of LDP in Binary Hypothesis Testing

Notation. Let $P^{\otimes n}$ be the distribution of

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} P.$$

Stein's Lemma. If $\varepsilon \in (0,1)$, then

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta(P^{\otimes n}, Q^{\otimes n})} = D_{\mathrm{KL}}(P\|Q).$$

Under (ε, δ)-LDP: $P \to KP$, $Q \to KQ$ &

$D_{\mathrm{KL}}(KP\|KQ) \quad \eta_{\mathrm{KL}}(K) D_{\mathrm{KL}}(P\|Q)$.

By the relation between f and ,

$$\eta_{\mathrm{KL}}(K) \quad 1 \quad \frac{1 \quad e^{\varepsilon}(K)}{e^{\varepsilon}}.$$

Since K is (ε, δ)-LDP , $e^{\varepsilon}(K)$,

$$\eta_{\mathrm{KL}}(K) \quad 1 \quad \frac{1}{e^{\varepsilon}}.$$

Thm. If K is (ε, δ)-LDP, then $\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta(KP^{\otimes n}, KQ^{\otimes n})} \quad 1 \quad \frac{1}{e^{\varepsilon}} \quad D_{\mathrm{KL}}(P\|Q)$.

Y. Polyanskiy & Y. Wu. Lecture Notes on Information Theory, 2014.
S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

Three Archetypal Problems in Privacy Analysis

Understand privacy from a `fundamental' perspective

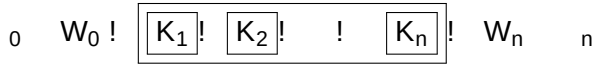
Assess the cost of privacy in speci c statistical applications

Compute/estimate the privacy parameters of speci c mechanisms

DP under Composition: Iterative Algorithms

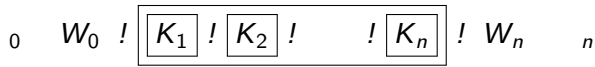
S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

DP under Composition: Iterative Algorithms



S. Asoodeh, M. Diaz & F. Calmon. \Contraction of E -divergence and its applications to privacy," arXiv:2012.11035.

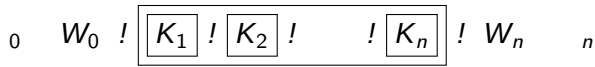
DP under Composition: Iterative Algorithms



DP for iterative algorithms: $\sup_{0;\ {}^{\theta}_{0} 2 P(X)} \mathrm{E}\ (\ {}_{n}k\ {}^{0}_{n})$ with $= e$

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.

DP under Composition: Iterative Algorithms



DP for iterative algorithms: $\sup_{0;\ {}^{\theta}_{0} 2P(X)} \mathrm{E}\ (\ {}_{n}k\ {}^{0}_{n})$ with $= e$

Contraction philosophy: $\mathrm{E}\ (\ {}_{n}k\ {}^{0}_{n})\quad \mathrm{E}\ (K_n)\quad \mathrm{E}\ (K_1)$

S. Asoodeh, M. Diaz & F. Calmon. "Contraction of E -divergence and its applications to privacy," *arXiv:2012.11035*.

Final Remarks

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

Archetypal Problems in Privacy Analysis

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

Archetypal Problems in Privacy Analysis

Understand privacy from a ‘fundamental’ perspective

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

Archetypal Problems in Privacy Analysis

Understand privacy from a 'fundamental' perspective

Quantify the cost of privacy in statistical applications

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

Archetypal Problems in Privacy Analysis

Understand privacy from a 'fundamental' perspective

Quantify the cost of privacy in statistical applications

Estimate the privacy parameters of specific mechanisms

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

Archetypal Problems in Privacy Analysis

Understand privacy from a 'fundamental' perspective

Quantify the cost of privacy in statistical applications

Estimate the privacy parameters of specific mechanisms

Provide statistical evidence of incorrect implementation

Slides available at *https://www.mariodiaztorres.com/*

Final Remarks

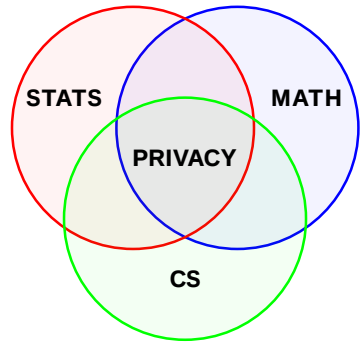
Archetypal Problems in Privacy Analysis

Understand privacy from a 'fundamental' perspective

Quantify the cost of privacy in statistical applications

Estimate the privacy parameters of specific mechanisms

Provide statistical evidence of incorrect implementation



Slides available at *https://www.mariodiaztorres.com/*