دانشگاه صنعتی شریف
دانشکده علوم ریاضی

پایان‌نامه کارشناسی ارشد
ریاضی کاربردی

تحلیل نظریه اطلاعاتیِ محرمانگی تفاضلی موضعی و کاربردهای آماری آن

نگارش

فیروزه ابریشمی

استاد راهنما

دکتر جواد ابراهیمی بروجنی

بهمن ۱۴۰۴

بسم الله الرحمن الرحیم

به نام خدا

دانشگاه صنعتی شریف

دانشکده علوم ریاضی

پایان‌نامه کارشناسی ارشد

این پایان‌نامه به عنوان تحقق بخشی از شرایط دریافت درجه کارشناسی ارشد است.

عنوان: تحلیل نظریه اطلاعاتیِ محرمانگی تفاضلی موضعی و کاربردهای آماری آن

نگارش: فیروزه ابریشمی

کمیته ممتحنین

استاد راهنما: دکتر جواد ابراهیمی بروجنی امضاء:

استاد مشاور: استاد مشاور امضاء:

استاد مدعو: استاد ممتحن امضاء:

تاریخ:

اظهارنامه

(اصالت متن و محتوای پایان‌نامه کارشناسی ارشد)

عنوان پایان‌نامه: تحلیل نظریه اطلاعاتیِ محرمانگی تفاضلی موضعی و کاربردهای آماری آن

استاد راهنما: دکتر جواد ابراهیمی بروجنی **استاد مشاور:** استاد مشاور

این‌جانب فیروزه ابریشمی اظهار می‌دارم:

۱. متن و نتایج علمی ارائه شده در این پایان‌نامه اصیل بوده و زیرنظر استادان نام‌برده شده در بالا تهیه شده است.

۲. متن پایان‌نامه به این صورت در هیچ جای دیگری منتشر نشده است.

۳. متن و نتایج مندرج در این پایان‌نامه، حاصل تحقیقات این‌جانب به عنوان دانشجوی کارشناسی ارشد دانشگاه صنعتی شریف است.

۴. کلیه مطالبی که از منابع دیگر در این پایان‌نامه مورد استفاده قرار گرفته، با ذکر مرجع مشخص شده است.

نگارنده: فیروزه ابریشمی
تاریخ:
امضاء:

استاد راهنما: دکتر جواد ابراهیمی بروجنی
تاریخ:
امضاء:

نگارنده: فیروزه ابریشمی
تاریخ:
امضاء:

چکیده

کلیدواژه‌ها:

فهرست مطالب

فهرست جداول

فهرست تصاویر

فصل ۱

مقدمه

۱-۱ اهمیت موضوع

۱-۲ ادبیات موضوع

۱-۳ اهداف پژوهش

۱-۴ ساختار پایان‌نامه

فصل ۲

پیش‌نیازها

۲-۱ محرمانگی تفاضلی متمرکز (CDP)

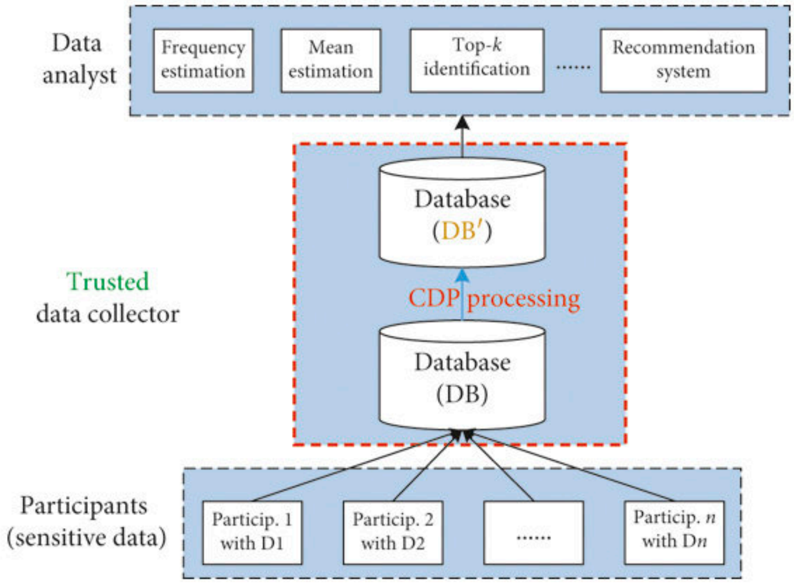
مفهوم محرمانگی تفاضلی[1] یا به اختصار ε-DP، اولین بار توسط دِوُرک و همکاران[؟] معرفی شد و به سرعت به استاندارد طلایی برای حفظ حریم خصوصی در تحلیل داده‌ها تبدیل گشت. این چارچوب، یک تعریف ریاضی قوی از حریم خصوصی ارائه می‌دهد که مبتنی بر پنهان‌سازی حضور یا عدم حضور یک فرد خاص در مجموعه داده است.

۲-۱-۱ مدل اعتماد و تعریف رسمی

در مدل متمرکز[2]، فرض بر این است که یک متصدی مورد اعتماد[3] وجود دارد. تمام افراد داده‌های خام و حساس خود را در اختیار این متصدی قرار می‌دهند (شکل ۲-۱ را ببینید). متصدی، مجموعه داده‌ی کامل $\mathcal{D}$ را در اختیار دارد. وظیفه‌ی متصدی این است که با اجرای یک مکانیزم تصادفی[4] $\mathcal{M}$ بر روی مجموعه داده‌ی $\mathcal{D}$، نتایجی (مثلاً پاسخ به یک پرس‌وجو[5]) را به صورت عمومی منتشر کند، به طوری که اطلاعات حساس افراد فاش نشود.

برای تعریف رسمی محرمانگی تفاضلی، مفاهیم **الگوریتم(مکانیزم) تصادفی**، **فاصله‌ی بین دو**

[1] Differential Privacy
[2] Centralized
[3] Trusted Curator
[4] Randomized Mechanism
[5] Query



شکل ۲-۱: مدل محرمانگی تفاضلی متمرکز با یک متصدی مورد اعتماد.

پایگاه‌داده[۶] و سپس **همسایگی**[۷] را تعریف می‌کنیم.

تعریف ۲-۱ (الگوریتم تصادفی) *طبق تعریف کتاب اضافه شود. عکسش تو همین فولدر هست. همچنین تعریف $\mathcal{X}$ و پایگاه‌داده هم اضافه شود.*

تعریف ۲-۲ (فاصله بین دو پایگاه‌داده) *نرم ℓ_1 یک پایگاه‌داده‌ی $\mathcal{D}$ به صورت $\|\mathcal{D}\|_1$ نمایش و به صورت زیر تعریف می‌شود:*

$$\|\mathcal{D}\|_1 = \sum_{i=1}^{|\mathcal{X}|} |x_i|$$

فاصله‌ی ℓ_1 بین دو پایگاه‌داده $\mathcal{D}_1$ و $\mathcal{D}_2$ برابر است با $\|\mathcal{D}_1 - \mathcal{D}_2\|_1$.

تعریف ۲-۳ (پایگاه‌داده‌های همسایه) *دو پایگاه‌داده‌ی $\mathcal{D}_1$ و $\mathcal{D}_2$ را همسایه*[۸] *می‌گوییم (و با $\mathcal{D}_1 \sim \mathcal{D}_2$ نشان می‌دهیم) اگر $\|\mathcal{D}_1 - \mathcal{D}_2\|_1 \leqslant 1$.*

ایده‌ی اصلی محرمانگی تفاضلی این است که خروجی مکانیزم برای دو مجموعه داده‌ی همسایه باید از نظر آماری «شبیه» باشد، به طوری که مهاجم نتواند تشخیص دهد ورودی واقعی کدام بوده است.

[6] Database
[7] Neighboring
[8] Adjacent

تعریف ۲-۴ (εــمحرمانگی تفاضلی **(ε-DP))** *یک مکانیزم تصادفی* $\mathcal{M}$*، تعریف* ε*ــ محرمانگی تفاضلی را برآورده می‌سازد، اگر برای هر دو مجموعه داده‌ی همسایه‌ی* $\mathcal{D}_۱$ *و* $\mathcal{D}_۲$ *و برای هر زیرمجموعه* $\mathcal{S}$ *از خروجی‌های ممکن* ($\mathrm{Range}(\mathcal{M})$)*، داشته باشیم:*

$$\Pr[\mathcal{M}(\mathcal{D}_۱) \in \mathcal{S}] \leqslant \exp(\varepsilon) \cdot \Pr[\mathcal{M}(\mathcal{D}_۲) \in \mathcal{S}] \quad (۱-۲)$$

گاهی اوقات، یک تعریف انعطاف‌پذیرتر به نام (ε, δ)-DP نیز استفاده می‌شود که اجازه‌ی یک احتمال شکست کوچک δ را می‌دهد:

$$\Pr[\mathcal{M}(\mathcal{D}_۱) \in \mathcal{S}] \leqslant \exp(\varepsilon) \cdot \Pr[\mathcal{M}(\mathcal{D}_۲) \in \mathcal{S}] + \delta \quad (۲-۲)$$

۲-۱-۲ مکانیزم‌های پایه در CDP

برای دستیابی به ε-DP، باید به پاسخ دقیق پرس‌وجو «نویز»[۹] اضافه کنیم. میزان نویز به حساسیت[۱۰] پرس‌وجو بستگی دارد.

تعریف ۲-۵ (حساسیت سراسری) *برای یک تابع* f*، حساسیت سراسری* $\ell_۱$ ($\Delta_۱ f$) *و* $\ell_۲$ ($\Delta_۲ f$) *به صورت زیر تعریف می‌شوند:*

$$\Delta_۱ f = \max_{\mathcal{D}_۱ \sim \mathcal{D}_۲} \|f(\mathcal{D}_۱) - f(\mathcal{D}_۲)\|_۱ \quad (۳-۲)$$

$$\Delta_۲ f = \max_{\mathcal{D}_۱ \sim \mathcal{D}_۲} \|f(\mathcal{D}_۱) - f(\mathcal{D}_۲)\|_۲ \quad (۴-۲)$$

سه مکانیزم اساسی برای دستیابی به CDP عبارتند از:

- **مکانیزم لاپلاس**[۱۱]: برای توابع عددی، با افزودن نویز از توزیع لاپلاس متناسب با حساسیت $\ell_۱$، می‌توان به ε-DP دست یافت:

$$\mathcal{M}(\mathcal{D}) = f(\mathcal{D}) + \mathrm{Lap}\left(\frac{\Delta_۱ f}{\varepsilon}\right) \quad (۵-۲)$$

- **مکانیزم گوسی**[۱۲]: این مکانیزم اغلب زمانی استفاده می‌شود که حساسیت $\ell_۲$ تابع کمتر از حساسیت $\ell_۱$ باشد. در اینجا نویز از توزیع نرمال (گوسی) افزوده می‌شود:

$$\mathcal{M}(\mathcal{D}) = f(\mathcal{D}) + \mathcal{N}\left(۰, \sigma^۲\right) \quad (۶-۲)$$

---
[9] Noise
[10] Sensitivity
[11] Laplace
[12] Gaussian

که در آن $\sigma \geqslant \sqrt{۲\ln(۱/۲۵/\delta)} \cdot \frac{\Delta_۲ f}{\varepsilon}$ است. برخلاف مکانیزم لاپلاس، این مکانیزم تنها (ε,δ)-DP را (با $\delta > ۰$) تضمین می‌کند.

- **مکانیزم نمایی**[۱۳]: برای خروجی‌های غیرعددی (دسته‌ای)، از یک «تابع امتیاز» $q(\mathcal{D}, r)$ استفاده می‌شود. این مکانیزم خروجی r را با احتمالی متناسب با امتیاز آن برمی‌گرداند:

$$\Pr[\mathcal{M}(\mathcal{D}) = r] \propto \exp\left(\frac{\varepsilon \cdot q(\mathcal{D}, r)}{۲\Delta q}\right) \tag{۲-۷}$$

۲-۱-۳ خواص کلیدی محرمانگی تفاضلی

قدرت چارچوب DP در خواص ترکیبی آن نهفته است:

- **مصونیت در برابر پس‌پردازش**[۱۴]: انجام هرگونه محاسبات بر روی خروجی یک مکانیزم ε-DP(بدون دسترسی مجدد به داده‌های اصلی)، نمی‌تواند سطح محرمانگی را کاهش دهد.

- **ترکیب‌پذیری**[۱۵]: اگر چندین مکانیزم ε-DPرا اجرا کنیم، بودجه‌های محرمانگی جمع می‌شوند.
  - *ترکیب‌پذیری پایه‌ای:* اجرای k مکانیزم ε_i-DP منجر به $\sum \varepsilon_i$-DP می‌شود.
  - *ترکیب‌پذیری پیشرفته:* با پذیرش یک δ کوچک، می‌توان نشان داد که بودجه کل با نرخ $\sqrt{k}$ رشد می‌کند (نه k).

- **محرمانگی گروهی**[۱۶]: محرمانگی تفاضلی به طور طبیعی برای گروه‌هایی از افراد نیز صادق است. اگر دو پایگاه داده در k رکورد با هم متفاوت باشند، تضمین محرمانگی به صورت $k\varepsilon$-DP برقرار خواهد بود. این یعنی با افزایش اندازه گروه، تضمین محرمانگی به صورت خطی تضعیف می‌شود.

۲-۱-۴ محدودیت مدل متمرکز

با وجود تمام مزایا، مدل CDP یک نقطه‌ی ضعف اساسی دارد: نیاز به یک متصدی کاملاً مورد اعتماد. در بسیاری از سناریوهای دنیای واقعی (مانند جمع‌آوری داده از گوشی‌های هوشمند)، کاربران به سرور مرکزی اعتماد ندارند. این عدم اعتماد، ما را به سمت مدل جایگزین، یعنی «محرمانگی تفاضلی موضعی» سوق می‌دهد. در بخش بعد با محرمانگی تفاضلی موضعی و تعاریف و قضایای اساسی آن آشنا خواهیم شد.

[13] Exponential
[14] Post-processing Immunity
[15] Composition
[16] Group Privacy

۲-۲ محرمانگی تفاضلی موضعی

در فصل گذشته، مبانی نظری حریم خصوصی متمرکز (CDP) و ابزارهای آماری لازم برای تحلیل آن را مرور کردیم. در این فصل، به طور اختصاصی به **چارچوب محرمانگی تفاضلی موضعی**[۱۷] **(α-LDP)** می‌پردازیم. این مدل، که امروزه در سیستم‌های توزیع‌شده و جمع‌آوری داده‌های بزرگ‌مقیاس کاربرد فراوان دارد، پارادایم اعتماد را از «سرور مرکزی» به «کاربر نهایی» تغییر می‌دهد.

۲-۲-۱ مقدمه و گذار از مدل متمرکز

همان‌طور که در بخش ۲-۱ دیدیم، مدل متمرکز نیازمند وجود یک متصدی مورد اعتماد[۱۸] است که به داده‌های خام دسترسی داشته باشد. اگرچه این مدل دقت آماری بالایی را فراهم می‌کند، اما در دنیای واقعی با چالش‌های امنیتی و حقوقی جدی روبروست:

- **نقطه شکست مرکزی**[۱۹]: سرور مرکزی هدف جذابی برای مهاجمان است. نشت اطلاعات از سرور (چه بر اثر هک و چه بر اثر خطای انسانی) حریم خصوصی تمام کاربران را به خطر می‌اندازد.
- **عدم اعتماد کاربران:** در بسیاری از کاربردها (مانند جمع‌آوری داده‌های پزشکی یا تاریخچه مرورگر)، کاربران تمایلی ندارند داده‌های حساس خود را حتی به یک سرور «مطمئن» بسپارند.

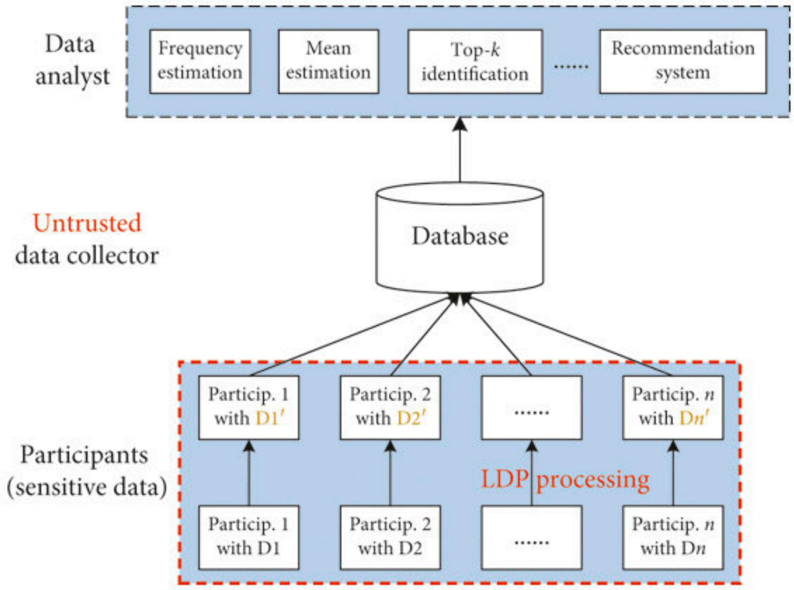
در پاسخ به این چالش‌ها، مدل محرمانگی تفاضلی موضعی مطرح شد. در α-LDP، فرآیند خصوصی‌سازی (افزودن نویز) به سمت کلاینت (کاربر) منتقل می‌شود. به این معنا که داده‌ها *قبل* از ترک دستگاه کاربر، نویزدار می‌شوند و سرور تنها به داده‌های بی‌نام و نویزدار دسترسی دارد (شکل ۲-۲).

این رویکرد توسط شرکت‌های بزرگ فناوری برای جمع‌آوری داده‌های تله‌متری پذیرفته شده است. برای مثال، گوگل از مکانیزم RAPPOR در مرورگر کروم، و اپل و مایکروسافت از روش‌های مشابهی برای جمع‌آوری داده‌های آماری از سیستم‌عامل‌های خود استفاده می‌کنند.

۲-۲-۲ تعاریف رسمی و مدل‌های محاسباتی

در مدل موضعی، ما n کاربر داریم که هر کدام یک داده‌ی خصوصی X_i از دامنه $\mathcal{X}$ در اختیار دارند. هر کاربر به طور مستقل یک الگوریتم تصادفی (مکانیزم) را اجرا می‌کند و خروجی Z_i را منتشر می‌کند.

[17] Local Differential Privacy (LDP)
[18] Trusted Curator
[19] Single Point of Failure



شکل ۲-۲: مدل محرمانگی تفاضلی موضعی (α-LDP). نویز به صورت موضعی روی دستگاه کاربر اضافه می‌شود.

تعریف α-LDP

هسته‌ی اصلی این مدل، «تصادفی‌ساز موضعی» است.

تعریف ۲-۶ (تصادفی‌ساز موضعی[۲۰]) *یک مکانیزم تصادفی $\mathcal{M} : \mathcal{X} \to \mathcal{Z}$ را یک تصادفی‌ساز موضعی می‌نامیم که ورودی $x \in \mathcal{X}$ را می‌گیرد و خروجی $z \in \mathcal{Z}$ را بر اساس توزیع احتمال شرطی $Q(z|x)$ تولید می‌کند.*

شرط محرمانگی در اینجا تضمین می‌کند که با مشاهده‌ی خروجی z، تمایز قائل شدن بین هر دو ورودی اولیه x و x' دشوار باشد. تفاوت کلیدی این تعریف با مدل متمرکز در این است که در CDP ما دو پایگاه داده‌ی همسایه را مقایسه می‌کردیم، اما در اینجا هر دو مقدار ورودی ممکن مقایسه می‌شوند.

تعریف ۲-۷ (αـ محرمانگی تفاضلی موضعی) *یک مکانیزم $\mathcal{M}$ دارای αـ محرمانگی تفاضلی موضعی (α-LDP) است اگر برای تمام جفت ورودی‌های $x, x' \in \mathcal{X}$ و هر زیرمجموعه از خروجی‌ها $\mathcal{S} \subseteq \mathcal{Z}$ داشته باشیم:*

$$\sup_{S} \frac{\Pr[\mathcal{M}(x) \in \mathcal{S}]}{\Pr[\mathcal{M}(x') \in \mathcal{S}]} \leqslant e^{\alpha} \qquad (۲-۸)$$

(نکته: در متون آماری مانند [؟] معمولاً از پارامتر α به جای ε برای نمایش بودجه حریم خصوصی موضعی استفاده می‌شود تا تمایز آن با مدل متمرکز مشخص باشد. ما نیز در این فصل و فصول بعدی از این نمادگذاری پیروی می‌کنیم).

این تعریف معادل شرط زیر بر روی واگرایی ماکزیمم (D_∞) بین توزیع‌های شرطی است:

$$\sup_{x,x' \in \mathcal{X}} D_\infty(Q(\cdot|x)||Q(\cdot|x')) \leqslant \alpha \tag{۲-۹}$$

تعمیم‌ها و خواص

علاوه بر تعریف استاندارد α-LDP (معادله ۲-۸)، دو مفهوم دیگر نیز در تحلیل‌های نظری و طراحی مکانیزم‌ها اهمیت دارند: محرمانگی تقریبی و خاصیت ترکیب.

تعریف ۲-۸ (((α,δ)ـ محرمانگی تفاضلی موضعی) *یک مکانیزم تصادفی $\mathcal{M}$ دارای محرمانگی تفاضلی موضعی تقریبی یا (α,δ)-LDP است اگر برای هر دو ورودی $x, x' \in \mathcal{X}$ و هر زیرمجموعه خروجی $\mathcal{S} \subseteq \mathcal{Z}$ داشته باشیم:*

$$\Pr[\mathcal{M}(x) \in \mathcal{S}] \leqslant e^\alpha \cdot \Pr[\mathcal{M}(x') \in \mathcal{S}] + \delta \tag{۲-۱۰}$$

این تعریف (که در [؟] نیز بررسی شده است)، اجازه‌ی یک احتمال شکست کوچک δ را می‌دهد. اهمیت نظری این تعریف در آن است که ارتباط مستقیمی با واگرایی E_γ (که در فصل قبل معرفی شد) دارد.

قضیه‌ی ۲-۱ (ترکیب ترتیبی[۲۱]) *اگر یک کاربر در k مرحله‌ی مختلف در پروتکل‌های $\mathcal{M}_1, \ldots, \mathcal{M}_k$ شرکت کند که هر کدام به ترتیب دارای بودجه‌ی حریم خصوصی α_i باشند، آنگاه کل فرآیند دارای محرمانگی تفاضلی موضعی با بودجه‌ی $\sum_{i=1}^k \alpha_i$ خواهد بود. این خاصیت در تحلیل پروتکل‌های تعاملی (بخش بعد) که در آن خروجی‌های بعدی به خروجی‌های قبلی وابسته هستند، نقش بنیادین دارد.*

پروتکل‌های تعاملی و غیرتعاملی

یکی از جنبه‌های مهم در تحلیل نرخ‌های مینیماکس (که در فصل بعد به آن می‌پردازیم)، نحوه‌ی تعامل کاربران با سرور است. دوچی و همکاران [؟] پروتکل‌های موضعی را به دو دسته تقسیم می‌کنند:

۱. **پروتکل‌های غیرتعاملی[۲۲]**: در این حالت، خروجی هر کاربر Z_i تنها به ورودی خودش X_i وابسته است و مستقل از داده‌ها یا خروجی‌های سایر کاربران تولید می‌شود.

$$Z_i = \mathcal{M}_i(X_i) \tag{۲-۱۱}$$

[22] Non-interactive

این مدل ساده‌ترین و رایج‌ترین شکل پیاده‌سازی α-LDPاست.

۲. پروتکل‌های تعاملی (ترتیبی)[۲۳]: در این حالت، مکانیزم کاربر i می‌تواند به خروجی‌های منتشر شده توسط کاربران قبلی $(Z_۱, \ldots, Z_{i-۱})$ وابسته باشد. به عبارت دیگر، کانال ارتباطی Q_i می‌تواند به صورت پویا بر اساس تاریخچه تغییر کند:

$$Z_i \sim Q_i(\cdot|X_i, Z_۱, \ldots, Z_{i-۱}) \quad (۲-۱۲)$$

این مدل به الگوریتم‌های تطبیقی اجازه می‌دهد تا دقت تخمین را بهبود بخشند. با این حال، همان‌طور که در فصل بعد خواهیم دید، حتی با وجود تعامل، محدودیت‌های بنیادی fـ واگرایی همچنان مانع از کاهش چشمگیر نرخ خطا می‌شوند.

۲-۲-۳ مکانیزم‌های پایه در α-LDP

در این بخش، مکانیزم‌های بنیادین را معرفی می‌کنیم که برای تحقق محرمانگی تفاضلی موضعی استفاده می‌شوند. این مکانیزم‌ها بلوک‌های سازنده‌ی پروتکل‌های پیچیده‌تر هستند و بسته به نوع داده (دودویی، دسته‌ای یا عددی) و اندازه دامنه انتخاب می‌شوند.

پاسخ تصادفی دودویی (RR)

پایه‌ای‌ترین و کلاسیک‌ترین مکانیزم در مدل موضعی، «پاسخ تصادفی»[۲۴] است که دهه‌ها پیش از تعریف رسمی α-LDPتوسط وارنر [؟] برای نظرسنجی‌های حساس معرفی شد. فرض کنید دامنه ورودی دودویی باشد ($\mathcal{X} = \{۰, ۱\}$).

مکانیزم $\mathcal{M}_{RR}$ با ورودی $x \in \{۰, ۱\}$، خروجی $z \in \{۰, ۱\}$ را طبق احتمالات زیر تولید می‌کند:

$$\Pr[z = x] = p, \quad \Pr[z \neq x] = ۱ - p \quad (۲-۱۳)$$

برای اینکه این مکانیزم شرط α-LDPرا برآورده کند، طبق تعریف ۲-۷ باید نسبت احتمالات حداکثر e^α باشد:

$$\frac{\Pr[z = ۱|x = ۱]}{\Pr[z = ۱|x = ۰]} = \frac{p}{۱ - p} \leqslant e^\alpha \quad (۲-۱۴)$$

بنابراین، برای دستیابی به کمترین خطا، پارامتر احتمال p به صورت زیر تنظیم می‌شود:

$$p = \frac{e^\alpha}{۱ + e^\alpha} \quad (۲-۱۵)$$

[23] Sequential/Interactive
[24] Randomized Response (RR)

در این حالت، واریانس تخمین‌گر حاصل از این مکانیزم برابر است با:

$$\mathrm{Var}[\hat{x}] = \frac{e^{\alpha}}{(e^{\alpha}-۱)^{۲}} \tag{۲-۱۶}$$

این رابطه نشان می‌دهد که برای αهای کوچک، واریانس به سرعت (تقریباً با نرخ $۱/\alpha^{۲}$) افزایش می‌یابد که نشان‌دهنده‌ی هزینه بالای حریم خصوصی در دقت تخمین است.

پاسخ تصادفی تعمیم‌یافته (GRR)

زمانی که دامنه ورودی شامل $k > ۲$ عنصر باشد ($\mathcal{X} = \{۱, \ldots, k\}$)، از نسخه تعمیم‌یافته پاسخ تصادفی[۲۵] استفاده می‌شود [؟]. این روش تعمیم مستقیم RR برای دامنه‌های گسسته است.

در این مکانیزم، برای ورودی x:

$$\Pr[\mathcal{M}(x) = z] = \begin{cases} p & \text{if } z = x \\ q & \text{if } z \neq x \end{cases} \tag{۲-۱۷}$$

از آنجا که مجموع احتمالات باید ۱ باشد، داریم $p + (k-۱)q = ۱$. همچنین شرط α-LDP ایجاب می‌کند که $\frac{p}{q} \leqslant e^{\alpha}$. با حل این دستگاه معادلات، مقادیر بهینه p و q به صورت زیر به دست می‌آیند:

$$p = \frac{e^{\alpha}}{e^{\alpha}+k-۱}, \quad q = \frac{۱}{e^{\alpha}+k-۱} \tag{۲-۱۸}$$

این مکانیزم برای دامنه‌های کوچک (k کوچک) بسیار کارآمد است و واریانس آن بهینه است. اما با افزایش k، دقت آن به شدت کاهش می‌یابد زیرا احتمال گزارش پاسخ صحیح (p) با افزایش k کاهش می‌یابد و به سمت صفر میل می‌کند. بنابراین برای دامنه‌های بزرگ (مانند لغات یک دیکشنری)، GRR گزینه مناسبی نیست.

مکانیزم‌های مبتنی بر کدگذاری یگانی (UE)

برای غلبه بر مشکل کاهش دقت GRR در دامنه‌های بزرگ، خانواده‌ای از مکانیزم‌ها تحت عنوان «کدگذاری یگانی»[۲۶] توسعه یافته‌اند. این رویکرد اساس پروتکل مشهور RAPPOR گوگل را تشکیل می‌دهد [؟].

در این روش، فرآیند خصوصی‌سازی طی دو مرحله انجام می‌شود:

---
[25] Generalized Randomized Response (GRR)
[26] Unary Encoding (UE)

۱. **کدگذاری (Encoding):** ورودی $x \in \{۱, \ldots, k\}$ به یک بردار بیتی v به طول k تبدیل می‌شود که تنها در موقعیت x برابر با ۱ و در سایر جاها ۰ است (One-hot encoding).

۲. **اختلال (Perturbation):** هر بیت این بردار به صورت مستقل با استفاده از یک مکانیزم باینری معکوس می‌شود.

اگر v_i بیت iـام بردار کدگذاری شده باشد، خروجی z_i به صورت زیر تولید می‌شود:

$$\Pr[z_i = ۱] = \begin{cases} p & \text{if } v_i = ۱ \\ q & \text{if } v_i = ۰ \end{cases} \tag{۲-۱۹}$$

بر اساس انتخاب مقادیر p و q، دو مکانیزم مهم در این خانواده تعریف می‌شوند:

- **کدگذاری یگانی متقارن (SUE):** که به آن Basic RAPPOR نیز گفته می‌شود. در این حالت، احتمالات تغییر بیت به گونه‌ای انتخاب می‌شوند که مکانیزم متقارن باشد ($p + q = ۱$). مقادیر بهینه عبارتند از:

$$p = \frac{e^{\alpha/۲}}{e^{\alpha/۲} + ۱}, \quad q = \frac{۱}{e^{\alpha/۲} + ۱} \tag{۲-۲۰}$$

مزیت SUE این است که واریانس تخمین برای هر آیتم مستقل از تعداد کل آیتم‌ها (k) است، اما به دلیل استفاده از نیمی از بودجه حریم خصوصی ($\alpha/۲$) برای هر بیت، خطای آن همچنان قابل توجه است.

- **کدگذاری یگانی بهینه (OUE):** وانگ و همکاران [؟] نشان دادند که برای تخمین فراوانی در دامنه‌های بزرگ، نیازی به متقارن بودن مکانیزم نیست. در روش OUE[۲۷]، پارامترها به گونه‌ای تنظیم می‌شوند که اطلاعات بیت‌های ۱ (سیگنال اصلی) با بیشترین دقت حفظ شود ($p = ۱/۲$) و نویز روی بیت‌های ۰ (که تعدادشان زیاد است) کنترل شود:

$$p = \frac{۱}{۲}, \quad q = \frac{۱}{e^{\alpha} + ۱} \tag{۲-۲۱}$$

تحلیل‌های نظری و تجربی نشان می‌دهند که OUE برای αهای متوسط و بزرگ، واریانس کمتری نسبت به GRR و SUE دارد و استاندارد فعلی برای جمع‌آوری داده‌های دسته‌ای بزرگ‌مقیاس است.

[27] Optimized Unary Encoding

مکانیزم لاپلاس موضعی

برای داده‌های عددی (مثلاً $x \in [-۱, ۱]$)، استفاده از مکانیزم لاپلاس که در مدل متمرکز محبوب است، در مدل موضعی نیز ممکن است اما با چالش‌هایی همراه است.

حساسیت سراسری (Δ) در مدل موضعی برابر با قطر دامنه است، زیرا هر دو ورودی $x, x' \in \mathcal{X}$ باید از نظر مکانیزم غیرقابل تمایز باشند. اگر دامنه ورودی $\mathcal{X} = [-۱, ۱]$ باشد، حساسیت برابر است با:

$$\Delta = \max_{x,x'} |x - x'| = |۱ - (-۱)| = ۲ \quad (۲۲-۲)$$

بنابراین، مکانیزم لاپلاس موضعی خروجی را به صورت زیر تولید می‌کند:

$$\mathcal{M}_{Lap}(x) = x + \eta, \quad \eta \sim \text{Lap}\left(\frac{۲}{\alpha}\right) \quad (۲۳-۲)$$

نکته مهمی که دوچی و همکاران [؟] به آن اشاره کرده‌اند این است که برخلاف مدل متمرکز، مکانیزم لاپلاس در مدل موضعی برای ابعاد بالا ($d > ۱$) **زیر-بهینه**[۲۸] است و نرخ خطای آن بدتر از مکانیزم‌های پیشرفته‌تر (مانند نمونه‌برداری هایپرکیوب) است که در فصل‌های آینده به آن‌ها اشاره خواهیم کرد.

۲-۲-۴ چالش سودمندی در مدل موضعی

بهای عدم اعتماد به سرور، کاهش شدید سودمندی[۲۹] آماری است. از آنجایی که نویز به داده‌ی هر فرد به صورت مستقل اضافه می‌شود، خطای تجمیعی در مدل α-LDPبسیار بیش‌تر از مدل متمرکز CDP است.

برای رسیدن به سطح دقت مشابه، مدل موضعی معمولاً به تعداد کاربران n بسیار بیشتری نیاز دارد. به طور کلی، در حالی که خطای مکانیزم‌های CDP اغلب با $O(۱/n)$ کاهش می‌یابد، خطای مکانیزم‌های α-LDPمعمولاً با $O(۱/\sqrt{n})$ کاهش می‌یابد. این کاهش در «اندازه نمونه مؤثر» یکی از موضوعات اصلی است که در فصل آینده با استفاده از ابزارهای f-واگرایی آن را اثبات خواهیم کرد.

۲-۳ f-واگرایی‌ها

در بخش‌های قبلی، ما مکانیزم‌های محرمانگی تفاضلی را به عنوان روش‌هایی برای ایجاد «شباهت آماری» بین خروجی‌های دو پایگاه داده‌ی همسایه معرفی کردیم. در این بخش، ما ابزار ریاضیاتی اصلی برای سنجش

[28] Sub-optimal
[29] Utility

این «شباهت» یا «فاصله» بین توزیع‌های احتمالی را معرفی می‌کنیم. این ابزار، خانواده‌ی fــ واگرایی‌ها[۳۰] است که بسیاری از معیارهای رایج فاصله‌ی آماری را به عنوان حالت‌های خاص خود در بر می‌گیرد.

۲ـ۳ـ۱ تعریف fــ واگرایی

مفهوم fــ واگرایی اولین بار توسط سیسر [؟] و به طور همزمان توسط علی و سیلوِی [؟] معرفی شد. این معیار، یک روش عمومی برای اندازه‌گیری تفاوت بین دو توزیع احتمال P و Q (تعریف شده بر روی یک فضای یکسان) ارائه می‌دهد.

تعریف ۲ـ۹ (fــ واگرایی) *فرض کنید P و Q دو توزیع احتمال باشند به طوری که P نسبت به Q مطلقاً پیوسته[۳۱] باشد. فرض کنید p و q توابع چگالی احتمال (یا توابع جرم احتمال) آن‌ها باشند. برای هر تابع محدب $f:(0,\infty)\to\mathbb{R}$ که $f(1)=0$ باشد، fــ واگرایی P از Q به صورت زیر تعریف می‌شود:*

$$D_f(P||Q) = \int q(x) f\left(\frac{p(x)}{q(x)}\right) dx \tag{۲ـ۲۴}$$

در حالت گسسته، این تعریف به صورت حاصل جمع زیر در می‌آید:

$$D_f(P||Q) = \sum_{x\in\mathcal{X}} q(x) f\left(\frac{p(x)}{q(x)}\right) \tag{۲ـ۲۵}$$

تابع f «ژنراتور» (تولیدکننده) واگرایی نامیده می‌شود و با انتخاب fهای متفاوت، می‌توان معیارهای فاصله یا واگرایی متفاوتی را به دست آورد. شرط $f(1)=0$ تضمین می‌کند که اگر دو توزیع یکسان باشند ($P=Q$)، واگرایی آن‌ها صفر خواهد بود.

۲ـ۳ـ۲ نمونه‌های مهم fــ واگرایی

بسیاری از معیارهای معروف در آمار و نظریه اطلاعات، حالت‌های خاصی از fــ واگرایی هستند:

- واگرایی کولبک‌ـ‌لایبلر[۳۲]: این معیار که به آن آنتروپی نسبی[۳۳] نیز گفته می‌شود، با انتخاب تابع $f(t)=t\log t$ به دست می‌آید:

$$D_{KL}(P||Q) = \sum_{x} p(x) \log\left(\frac{p(x)}{q(x)}\right) \tag{۲ـ۲۶}$$

[30] f-divergences
[31] Absolutely Continuous
[32] Kullback-Leibler (KL) Divergence
[33] Relative Entropy

- فاصله‌ی واریانس کل[۳۴]: فاصله‌ی TV (که اغلب با $\Delta(P,Q)$ یا d_{TV} نشان داده می‌شود) ارتباط نزدیکی با fـ واگرایی دارد و با انتخاب $f(t) = \frac{۱}{۲}|t-۱|$ حاصل می‌شود:

$$d_{TV}(P,Q) = \frac{۱}{۲}\sum_x |p(x) - q(x)| \tag{۲-۲۷}$$

- واگرایی کای‌ـ دو[۳۵]: این معیار آماری با انتخاب $f(t) = (t-۱)^۲$ به دست می‌آید:

$$\chi^۲(P||Q) = \sum_x \frac{(p(x)-q(x))^۲}{q(x)} \tag{۲-۲۸}$$

- فاصله‌ی هلینجر (مربع)[۳۶]: این فاصله با انتخاب $f(t) = (\sqrt{t}-۱)^۲$ (یا معادل آن $f(t) = \frac{۱}{۲}(\sqrt{t}-۱)^۲$) حاصل می‌شود:

$$H^۲(P,Q) = \sum_x (\sqrt{p(x)} - \sqrt{q(x)})^۲ \tag{۲-۲۹}$$

۲-۳-۳ ارتباط fـ واگرایی‌ها با یکدیگر

این واگرایی‌ها مستقل از هم نیستند و روابط ریاضی مهمی بین آن‌ها برقرار است. یکی از مشهورترین این روابط، نامساوی پینسکر[۳۷] است که ارتباط بین واگرایی KL و فاصله‌ی TV را نشان می‌دهد:

$$d_{TV}(P,Q)^۲ \leqslant \frac{۱}{۲} D_{KL}(P||Q) \tag{۲-۳۰}$$

این نامساوی‌ها در تحلیل‌های حریم خصوصی بسیار کاربردی هستند، زیرا به ما اجازه می‌دهند که با داشتن یک کران (حد) بر روی یک معیار واگرایی، بتوانیم کرانی برای سایر معیارها نیز به دست آوریم.

در فصل بعدی، ما به تفصیل بررسی خواهیم کرد که چگونه تضمین α-LDP (که در معادله ۲-۸ تعریف شد) مستقیماً منجر به ایجاد یک کران بالا بر روی fـ واگرایی‌های مختلف بین توزیع‌های خروجی می‌شود.

۲-۴ مبانی آماری و نظریه اطلاعات

در بخش‌های پیشین، ابزارهای سنجش فاصله بین توزیع‌ها (مانند fـ واگرایی‌ها) را معرفی کردیم. در این بخش، به معرفی چارچوب آماری می‌پردازیم که در آن از این ابزارها برای تحلیل حدود پایین خطا در حضور

[34] Total Variation (TV) Distance
[35] Chi-Squared (χ^2) Divergence
[36] Squared-Hellinger Distance
[37] Pinsker's Inequality

محدودیت‌های محرمانگی استفاده می‌شود. این تعاریف و قضایا عمدتاً بر اساس چارچوب ارائه‌شده در [؟] تدوین شده‌اند.

۲-۴-۱ معیارهای فاصله اطلاعاتی

برای دو توزیع احتمال P و Q که روی فضای $\mathcal{X}$ تعریف شده‌اند و نسبت به یک اندازه‌ی پایه μ مطلقاً پیوسته هستند (با توابع چگالی p و q)، معیارهای زیر را تعریف می‌کنیم:

تعریف ۲-۱۰ (واگرایی کولبک-لایبلر) *واگرایی کولبک-لایبلر (KL) بین دو توزیع P و Q به صورت زیر تعریف می‌شود:*

$$D_{KL}(P||Q) = \int_{\mathcal{X}} p(x) \log \frac{p(x)}{q(x)} d\mu(x) \tag{۲-۳۱}$$

تعریف ۲-۱۱ (فاصله‌ی واریانس کل) *فاصله‌ی واریانس کل[۳۸] بین دو توزیع P و Q به صورت زیر تعریف می‌شود:*

$$||P - Q||_{TV} = \sup_{S \in \sigma(\mathcal{X})} |P(S) - Q(S)| = \frac{1}{2} \int_{\mathcal{X}} |p(x) - q(x)| d\mu(x) \tag{۲-۳۲}$$

تعریف ۲-۱۲ (اطلاعات متقابل) *اگر X و V دو متغیر تصادفی باشند، اطلاعات متقابل[۳۹] بین آن‌ها به صورت امید ریاضی واگرایی KL بین توزیع شرطی و توزیع حاشیه‌ای تعریف می‌شود:*

$$I(X; V) = D_{KL}(P_{X,V}||P_X \otimes P_V) = \mathbb{E}_V \left[D_{KL}(P_{X|V}||P_X)\right] \tag{۲-۳۳}$$

این معیار نقش کلیدی در نامساوی فانو (که در ادامه می‌آید) ایفا می‌کند.

۲-۴-۲ ریسک مینیماکس

در نظریه تصمیم آماری، هدف تخمین یک پارامتر $\theta(P)$ از یک توزیع ناشناخته $P \in \mathcal{P}$ است. اگر $\hat{\theta}$ یک تخمین‌گر باشد که تابعی از داده‌های مشاهده شده (مانند $Z_1, \dots, Z_n$) است، کیفیت آن با استفاده از یک تابع زیان صعودی $\Phi \circ \rho$ سنجیده می‌شود (که ρ یک شبه‌متر روی فضای پارامتر است).

نرخ مینیماکس[۴۰]، کمترین خطای ممکنی است که یک تخمین‌گر در بدترین سناریو (بدترین توزیع P در کلاس $\mathcal{P}$) متحمل می‌شود.

---
[38] Total Variation Distance
[39] Mutual Information
[40] Minimax Rate

تعریف ۲-۱۳ (نرخ مینیماکس) *برای یک کلاس از توزیع‌ها $\mathcal{P}$ و پارامتر θ، نرخ مینیماکس $\mathfrak{M}_n$ به صورت زیر تعریف می‌شود:*

$$\mathfrak{M}_n(\theta(\mathcal{P}), \Phi \circ \rho) = \inf_{\hat{\theta}} \sup_{P \in \mathcal{P}} \mathbb{E}_P[\Phi(\rho(\hat{\theta}(Z^n), \theta(P)))] \quad (۲-۳۴)$$

که در آن اینفیمم روی تمام تخمین‌گرهای ممکن $\hat{\theta}$ گرفته می‌شود.

در حالتی که محدودیت محرمانگی تفاضلی موضعی با پارامتر α وجود داشته باشد، نرخ مینیماکس خصوصی (α-Private Minimax Rate) با در نظر گرفتن اینفیمم روی تمام مکانیزم‌های کانال Q که شرط α-LDPرا برآورده می‌کنند، تعریف می‌شود [؟].

۲-۵ آزمون فرض آماری و روش تقلیل

برای اثبات حدود پایین نرخ‌های مینیماکس، روش استاندارد این است که مسئله‌ی تخمین پارامتر را به یک مسئله‌ی آزمون فرض[۴۱] تقلیل دهیم. ایده اصلی این است: اگر نتوانیم بین چند مقدار گسسته از پارامتر با دقت بالا تمایز قائل شویم، قطعاً نمی‌توانیم پارامتر را در فضای پیوسته با خطای کم تخمین بزنیم.

۲-۵-۱ آزمون فرض دودویی

ساده‌ترین حالت آزمون فرض، تصمیم‌گیری بین دو توزیع احتمال $P_۰$ و $P_۱$ است. فرض کنید داده‌ی مشاهده شده Z از یکی از این دو توزیع تولید شده است. ما دو فرض داریم:

- فرض صفر ($H_۰$): $Z \sim P_۰$
- فرض مقابل ($H_۱$): $Z \sim P_۱$

یک آزمون (یا تابع تست) $\psi : \mathcal{Z} \to \{۰, ۱\}$ تابعی است که بر اساس داده‌ی مشاهده شده، حدس می‌زند کدام فرض صحیح است. خطای این آزمون به صورت مجموع احتمال خطای نوع اول و دوم تعریف می‌شود:

$$P_{err}(\psi) = \Pr_{H_۰}(\psi(Z) = ۱) + \Pr_{H_۱}(\psi(Z) = ۰) \quad (۲-۳۵)$$

[41] Hypothesis Testing

لم نیمن–پیرسون[42] نشان می‌دهد که کمترین خطای ممکن برای هر آزمون دودویی، مستقیماً با فاصله‌ی واریانس کل (d_{TV}) بین دو توزیع ارتباط دارد:

$$\inf_{\psi} P_{err}(\psi) = ۱ - \|P_۰ - P_۱\|_{TV} \tag{۲-۳۶}$$

این رابطه نشان می‌دهد که هرچه دو توزیع $P_۰$ و $P_۱$ به هم شبیه‌تر باشند (فاصله‌ی TV کمتر)، احتمال خطا بیشتر شده و به ۱ (حدس تصادفی) نزدیک‌تر می‌شود. در فضای α-LDP، نویز اضافه شده باعث کاهش شدید فاصله‌ی TV و در نتیجه افزایش خطای آزمون می‌شود.

۲–۵–۲ تقلیل تخمین به آزمون (روش بسته‌بندی)

برای استفاده از ابزارهای آزمون فرض در مسئله‌ی تخمین نرخ مینیماکس (معادله ۲–۳۴)، از تکنیک گسسته‌سازی فضای پارامتر Θ استفاده می‌کنیم. این روش شامل مراحل زیر است:

۱. **ساخت مجموعه‌ی بسته‌بندی[43]:** مجموعه‌ای متناهی از پارامترها $\mathcal{V} = \{\theta_۱, \dots, \theta_M\} \subset \Theta$ را انتخاب می‌کنیم به طوری که از یکدیگر فاصله‌ی معناداری داشته باشند. به طور دقیق‌تر، اگر ρ متریک خطا باشد، برای هر $i \neq j$ باید داشته باشیم $\rho(\theta_i, \theta_j) \geqslant ۲\delta$.

۲. **تعریف مسئله‌ی آزمون:** فرض می‌کنیم طبیعت[44] یک اندیس V را به صورت تصادفی و یکنواخت از مجموعه $\{۱, \dots, M\}$ انتخاب می‌کند و داده‌ها بر اساس توزیع P_{θ_V} تولید می‌شوند. هدف، یافتن V بر اساس داده‌های مشاهده شده است.

۳. **ارتباط خطاها:** اگر یک تخمین‌گر $\hat{\theta}$ وجود داشته باشد که خطای تخمین آن با احتمال بالا کمتر از δ باشد، می‌توانیم از آن برای حل مسئله‌ی آزمون فرض استفاده کنیم (با انتخاب نزدیک‌ترین θ_i به $\hat{\theta}$). بنابراین، کران پایین روی خطای آزمون فرض، یک کران پایین برای خطای تخمین ایجاد می‌کند:

$$\mathfrak{M}_n(\theta(\mathcal{P})) \geqslant \Phi(\delta) \cdot \inf_{\psi} \Pr(\psi(Z^n) \neq V) \tag{۲-۳۷}$$

۲–۵–۳ نامساوی‌های کران پایین

برای اثبات کران‌های پایین، سه روش اصلی که بر پایه f–واگرایی‌ها بنا شده‌اند را معرفی می‌کنیم:

[42]Neyman-Pearson Lemma
[43]Packing Set
[44]Nature

قضیه‌ی ۲-۲ (نامساوی لو کم[۴۵]**)** *این روش برای آزمون بین دو توزیع $P_۱$ و $P_۲$ استفاده می‌شود. کمینه احتمال خطا با استفاده از فاصله‌ی واریانس کل (رابطه ۲-۳۲) کران‌دار می‌شود:*

$$\inf_{\psi} \Pr(\psi(Z^n) \neq V) \geqslant \frac{۱}{۲}\left(۱ - \|P_۱^n - P_۲^n\|_{TV}\right) \tag{۲-۳۸}$$

این روش زمانی مفید است که مسئله را به تشخیص بین دو حالت ساده تقلیل دهیم.

قضیه‌ی ۲-۳ (نامساوی فانو[۴۶]**)** *زمانی که پارامتر مورد نظر متعلق به مجموعه‌ای بزرگتر $\mathcal{V}$ باشد (تعداد فرضیه‌ها $|\mathcal{V}| > ۲$)، نامساوی فانو کران پایین قوی‌تری ارائه می‌دهد که مبتنی بر اطلاعات متقابل است:*

$$\inf_{\psi} \Pr(\psi(Z^n) \neq V) \geqslant ۱ - \frac{I(Z^n; V) + \log ۲}{\log |\mathcal{V}|} \tag{۲-۳۹}$$

که در آن V متغیر تصادفی یکنواخت روی مجموعه اندیس‌ها $\mathcal{V}$ است.

لم ۲-۴ (لم اسود[۴۷]**)** *این لم مسئله تخمین را به چندین آزمون فرض دودویی مستقل روی مختصات یک ابرمکعب $\{-۱, ۱\}^d$ تبدیل می‌کند. نسخه دقیق‌تر آن که در [؟] استفاده شده است، کران پایین را بر اساس فاصله‌ی واریانس کل توزیع‌های مخلوط حاشیه‌ای بیان می‌کند:*

$$\mathfrak{M}_n(\theta(\mathcal{P})) \geqslant \delta \sum_{j=۱}^{d} \left[۱ - \|M_{+j}^n - M_{-j}^n\|_{TV}\right] \tag{۲-۴۰}$$

که در آن M_{+j}^n و M_{-j}^n توزیع‌های حاشیه‌ای مخلوط روی مقادیر ۱+ و ۱− در بُعد jـ‌ام هستند.

فصل ۳

تحلیل‌های مبتنی بر انقباض و نرخ‌های مینیماکس

۳-۱ مقدمه

در فصل پیشین، تعاریف پایه محرمانگی تفاضلی موضعی (LDP) و مکانیزم‌های ابتدایی آن را بررسی کردیم. همان‌طور که دیدیم، چالش اصلی در مدل موضعی، کاهش شدید نسبت سیگنال به نویز است. برای تحلیل دقیق این پدیده و یافتن حدود نهایی دقت آماری، نیازمند ابزارهای قوی‌تری هستیم.

در این فصل، به بررسی چارچوب نظری استانداردی می‌پردازیم که توسط دوچی و همکاران [؟] توسعه داده شده است. ایده مرکزی این چارچوب، نگاه به مکانیزم‌های محرمانگی به عنوان «عملگرهای انقباضی»[1] است. به بیان شهودی، اعمال شرط α-LDP باعث می‌شود که توزیع‌های خروجی $\mathcal{M}(\cdot|x)$ و $\mathcal{M}(\cdot|x')$ بسیار به یکدیگر شبیه شوند، حتی اگر ورودی‌های x و x' کاملاً متفاوت باشند.

ما نشان خواهیم داد که چگونه می‌توان این شباهت اجباری را با استفاده از نامساوی‌های پردازش داده و fـ واگرایی‌ها (به‌ویژه واگرایی کولبک ـ لایبلر) مدل‌سازی کرد و از آن برای اثبات نرخ‌های مینیماکس در مسائل تخمین آماری استفاده نمود [؟].

۳-۲ محرمانگی به عنوان انقباض اطلاعاتی

یکی از ویژگی‌های بنیادین نظریه اطلاعات، «نامساوی پردازش داده»[2] است که بیان می‌کند پردازش روی داده‌ها (بدون دسترسی به منبع اصلی) نمی‌تواند اطلاعات متقابل را افزایش دهد. در زمینه محرمانگی، ما

[1] Contraction Operators
[2] Data Processing Inequality

با نسخه قوی‌تری از این مفهوم سروکار داریم که به آن «نامساوی قوی پردازش داده»[3] می‌گویند [؟].

فرض کنید $\mathcal{M}$ یک مکانیزم α-LDP باشد. هدف ما یافتن کرانی برای واگرایی بین توزیع‌های خروجی بر حسب واگرایی ورودی‌هاست. دوچی و همکاران نشان دادند که مکانیزم‌های موضعی باعث انقباض شدید در واگرایی KL می‌شوند.

قضیه‌ی ۳-۱ (انقباض KL در مکانیزم‌های موضعی) *فرض کنید $\mathcal{M}$ یک مکانیزم α-LDP باشد. برای هر دو ورودی $x, x' \in \mathcal{X}$، واگرایی کولبک-لایبلر بین توزیع‌های خروجی متناظر $\mathcal{M}(\cdot|x)$ و $\mathcal{M}(\cdot|x')$ با رابطه زیر محدود می‌شود:*

$$D_{KL}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) \leqslant ۴(e^{\alpha} - ۱)^{۲} \tag{۱-۳}$$

به طور دقیق‌تر، اگر $\alpha \leqslant ۱$ باشد، این کران به صورت $O(\alpha^{۲})$ رفتار می‌کند [؟].

اثبات. برای اثبات دقیق این قضیه، از تعریف واگرایی KL شروع می‌کنیم. فرض کنید $q(z|x)$ و $q(z|x')$ چگالی‌های احتمال خروجی باشند. طبق تعریف α-LDP می‌دانیم که برای هر $z \in \mathcal{Z}$:

$$e^{-\alpha} \leqslant \frac{q(z|x)}{q(z|x')} \leqslant e^{\alpha} \tag{۲-۳}$$

این شرط تضمین می‌کند که نسبت درست‌نمایی‌ها حول عدد ۱ محدود است. با بسط تیلور تابع $\log t$ حول $t = ۱$ و استفاده از خواص تحدب، می‌توان نشان داد که:

$$D_{KL}(P||Q) = \int p(z) \log \frac{p(z)}{q(z)} dz \tag{۳-۳}$$

$$\leqslant \int p(z) \left(\frac{p(z)}{q(z)} - ۱ + \frac{۱}{۲} \left(\frac{p(z)}{q(z)} - ۱ \right)^{۲} \right) dz \tag{۴-۳}$$

با اعمال کران‌های α-LDP بر روی نسبت p/q، جمله درجه اول صفر می‌شود و جمله درجه دوم ضریب $(e^{\alpha} - ۱)^{۲}$ را تولید می‌کند. جزئیات کامل این محاسبات در لم ۱ مقاله [؟] آمده است. نکته کلیدی این است که برای α کوچک، فاصله KL به صورت مربعی با α کاهش می‌یابد. □

این قضیه ابزار بسیار قدرتمندی است. به جای اینکه مستقیماً با تعریف دشوار α-LDP کار کنیم، می‌توانیم از این کران ساده در نامساوی‌هایی مثل فانو استفاده کنیم. همچنین مطالعات جدیدتر نشان داده‌اند که این انقباض را می‌توان با استفاده از معیارهای دیگری نظیر اطلاعات متقابل [؟] یا واگرایی E_{γ} [؟] نیز بیان کرد که در فصل بعد به آن می‌پردازیم.

---
[3] Strong Data Processing Inequality (SDPI)

۳-۲-۱ انقباض در فاصله واریانس کل

علاوه بر KL، کران مشابهی برای فاصله واریانس کل (TV) نیز ارائه شده است که در استفاده از روش «لم لو کم[۴]» کاربرد دارد [؟]:

قضیه‌ی ۳-۲ (انقباض TV) *تحت شرایط مشابه، برای هر مکانیزم α-LDP:*

$$\|\mathcal{M}(\cdot|x) - \mathcal{M}(\cdot|x')\|_{TV} \leqslant \min\{۱, e^{\alpha} - ۱\} \cdot \|x - x'\|. \qquad (۳-۵)$$

(در اینجا $\|x - x'\|$ نشان‌دهنده فاصله همینگ یا متریک مجزا روی ورودی است). برای α کوچک، این رابطه بیان می‌کند که فاصله آماری خروجی‌ها نمی‌تواند بیشتر از $O(\alpha)$ باشد.

۳-۳ تحلیل نرخ‌های مینیماکس با استفاده از انقباض

حال که ابزار انقباض را در اختیار داریم، می‌توانیم استراتژی کلی اثبات حدود پایین[۵] در مدل موضعی را صورت‌بندی کنیم. این استراتژی که توسط دوچی [؟] و بعدها با جزئیات بیشتر در [؟] بسط داده شد، شامل سه گام است:

۱. **تقلیل به آزمون فرض:** تبدیل مسئله تخمین پارامتر θ به مسئله تشخیص اندیس V در یک مجموعه متناهی (استفاده از لم فانو یا اسود).

۲. **کران‌دار کردن اطلاعات متقابل:** استفاده از خاصیت انقباض α-LDP برای محدود کردن اطلاعاتی که نمونه‌های مشاهده شده $Z_۱, \ldots, Z_n$ درباره اندیس V می‌دهند [؟].

۳. **محاسبه ریسک نهایی:** ترکیب نتایج برای رسیدن به کران پایین خطای تخمین.

مهم‌ترین گام، گام دوم است. طبق نامساوی قوی پردازش داده برای مدل موضعی، داریم:

$$I(V; Z^n) \leqslant \sum_{i=۱}^{n} I(V; Z_i) \leqslant n \cdot \alpha^۲ \cdot C \qquad (۳-۶)$$

که در آن C ثابتی است که به هندسه مسئله بستگی دارد. این رابطه نشان می‌دهد که اطلاعات موثر با نرخ $n\alpha^۲$ رشد می‌کند، نه n. این همان دلیلی است که «اندازه نمونه موثر» در مدل موضعی برابر با $n\alpha^۲$ در نظر گرفته می‌شود.

[4]Le Cam
[5]Lower Bounds

۳-۴ مطالعه موردی: تخمین میانگین

برای نمایش قدرت این چارچوب، مسئله کلاسیک تخمین میانگین را در نظر می‌گیریم. فرض کنید هر کاربر i برداری $X_i \in [-۱, ۱]^d$ دارد و هدف تخمین میانگین جامعه $\mu = \mathbb{E}_X$ است. معیار خطا را «میانگین مربعات خطا» (MSE) در نظر می‌گیریم.

قضیه‌ی ۳-۳ (کران پایین تخمین میانگین) *برای هر مکانیزم α-LDP و هر تخمین‌گر $\hat{\mu}$، ماکسیمم خطای مورد انتظار با رابطه زیر محدود می‌شود [؟]:*

$$\inf_{\hat{\mu},\mathcal{M}} \sup_{P} \mathbb{E}_{\|\hat{\mu}-\mu\|^۲} \geqslant \Omega\left(\frac{d}{n\min\{\alpha,\alpha^۲\}}\right) \qquad (۳-۷)$$

تحلیل اثبات: برای اثبات این کران، از لم اسود استفاده می‌کنیم. فضای پارامتر را به صورت یک ابرمکعب $\{-۱, ۱\}^d$ گسسته‌سازی می‌کنیم. طبق لم اسود، خطا با مجموع فاصله‌های TV بین توزیع‌های شرطی مرتبط است. با استفاده از قضیه انقباض ۳-۱ و نامساوی پینسکر، می‌دانیم که:

$$\|\mathcal{M}(\cdot|x) - \mathcal{M}(\cdot|x')\|_{TV}^۲ \leqslant \frac{۱}{۲} D_{KL}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) \leqslant O(\alpha^۲) \qquad (۳-۸)$$

بنابراین فاصله TV حداکثر از مرتبه α است. با جایگذاری این مقدار در لم اسود، کران پایین $\frac{۱}{n\alpha^۲}$ حاصل می‌شود.

این نتیجه نشان می‌دهد که برای رسیدن به خطای کم در مدل موضعی، تعداد داده‌ها باید متناسب با $۱/\alpha^۲$ افزایش یابد، که هزینه‌ی بسیار سنگین‌تری نسبت به مدل متمرکز (که متناسب با $۱/\varepsilon$ است) دارد.

۳-۵ محدودیت‌های تحلیل کلاسیک

با وجود موفقیت چارچوب دوچی در اثبات نرخ‌های مینیماکس بهینه برای αهای کوچک (رژیم محرمانگی بالا)، این روش در رژیم αهای بزرگ (محرمانگی پایین) دچار ضعف است.

همان‌طور که در رابطه (۳-۱) دیدیم، کران انقباض KL با ضریب $(e^\alpha - ۱)^۲$ رشد می‌کند. زمانی که α بزرگ باشد، این کران به سرعت به بی‌نهایت میل می‌کند و اطلاعاتی فراتر از کران بدیهی به ما نمی‌دهد. این در حالی است که به طور شهودی، حتی با α بزرگ، مکانیزم همچنان باید مقداری انقباض ایجاد کند.

این محدودیت ناشی از ذات واگرایی KL است که رفتار دنباله‌های توزیع را با حساسیت زیادی وزن‌دهی می‌کند. برای رفع این مشکل و به دست آوردن تحلیل‌های دقیق‌تر[۶] که در تمام بازه‌های α معتبر باشند، نیازمند

[6] Tight

معیار هندسی متفاوتی هستیم. این نیاز، انگیزه اصلی معرفی واگرایی‌های جدید مانند fـ واگرایی‌های خاص (نظیر E_γ) است [؟، ؟] که در فصل آینده به تفصیل به آن خواهیم پرداخت.

فصل ۴

هم‌ارزی α-LDP و انقباض E_γـ واگرایی

۴ـ۱ مقدمه و انگیزه

در فصل پیشین، دیدیم که چگونه دوچی و همکاران [؟] از واگرایی کولبک‌ـ لایبلر (KL) برای تحلیل محرمانگی تفاضلی موضعی استفاده کردند. اگرچه کران‌های آن‌ها برای رژیم‌های محرمانگی بالا (α کوچک) بسیار کارآمد هستند، اما در رژیم‌های α متوسط و بزرگ، دقت خود را از دست می‌دهند.

مشکل اصلی در آنجاست که واگرایی KL متریک «بومی» برای تعریف α-LDPنیست. تعریف α-LDP(معادله ۲ـ۸) مبتنی بر نسبت احتمالات است، در حالی که KL مبتنی بر لگاریتم نسبت‌هاست. این ناهمخوانی باعث می‌شود که در تبدیل شرایط α-LDPبه کران‌های KL، اطلاعاتی از دست برود (lossy conversion).

در این فصل، نشان می‌دهیم که یک معیار واگرایی دیگر به نام E_γـ واگرایی وجود دارد که دقیقاً ساختار هندسی α-LDPرا تسخیر می‌کند. ما ثابت خواهیم کرد که شرط α-LDPدقیقاً معادل صفر شدن E_γـ واگرایی (برای $\gamma = e^\alpha$) است. سپس از این هم‌ارزی برای استخراج کران‌های انقباض دقیق[1] برای سایر واگرایی‌ها استفاده خواهیم کرد که نتایج دوچی را بهبود می‌بخشند [؟].

[1] Tight

۴-۲ معرفی E_γـ واگرایی

E_γـ واگرایی یکی از اعضای کمتر شناخته‌شده‌ی خانواده fـ واگرایی‌هاست که در نظریه اطلاعات برای مقایسه نسبت درست‌نمایی توزیع‌ها کاربرد دارد.

تعریف ۴-۱ (E_γـ واگرایی) *فرض کنید P و Q دو توزیع احتمال باشند و $\gamma \geqslant 1$ یک عدد حقیقی باشد. E_γـ واگرایی بین P و Q به صورت زیر تعریف می‌شود:*

$$E_\gamma(P||Q) = \sup_{\mathcal{S}\in\sigma(\mathcal{X})} (P(\mathcal{S}) - \gamma Q(\mathcal{S})) \tag{۴-۱}$$

این تعریف را می‌توان به صورت بسته‌ی زیر نیز نوشت:

$$E_\gamma(P||Q) = \int_{\mathcal{X}} \max\{0, p(x) - \gamma q(x)\} d\mu(x) \tag{۴-۲}$$

که در آن p و q توابع چگالی احتمال هستند.

۴-۲-۱ خواص هندسی

این واگرایی خواص جالبی دارد که آن را برای تحلیل محرمانگی ایده‌آل می‌کند:

- **ارتباط با فاصله واریانس کل:** اگر $\gamma = 1$ باشد، داریم:

$$E_1(P||Q) = \sup_{\mathcal{S}}(P(\mathcal{S}) - Q(\mathcal{S})) = \|P - Q\|_{TV} \tag{۴-۳}$$

بنابراین E_γ تعمیمی از فاصله TV است.

- **غیرمنفی بودن:** همواره $E_\gamma(P||Q) \geqslant 0$ نیست. در واقع، اگر نسبت $p(x)/q(x)$ همواره کمتر از γ باشد، این مقدار صفر می‌شود. دقیقاً همین ویژگی است که آن را به α-LDPمرتبط می‌کند.

۴-۳ قضیه هم‌ارزی اصلی

اکنون به مهم‌ترین نتیجه‌ی نظری این پایان‌نامه می‌رسیم: اثبات اینکه α-LDPچیزی جز محدودیت بر روی E_γـ واگرایی نیست.

قضیه‌ی ۴-۱ (هم‌ارزی α-LDP و E_γ) *یک مکانیزم $\mathcal{M}$ در شرط α-LDP صدق می‌کند اگر و تنها اگر برای تمام جفت ورودی‌های $x, x' \in \mathcal{X}$:*

$$E_{e^\alpha}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) = ۰ \qquad (۴-۴)$$

اثبات. اثبات را در دو جهت انجام می‌دهیم.

جهت اول ($\Rightarrow$): فرض کنید $\mathcal{M}$ خاصیت α-LDP دارد. طبق تعریف ۲-۷، برای هر زیرمجموعه خروجی $\mathcal{S} \subseteq \mathcal{Z}$ و هر x, x' داریم:

$$\Pr[\mathcal{M}(x) \in \mathcal{S}] \leqslant e^\alpha \Pr[\mathcal{M}(x') \in \mathcal{S}] \qquad (۵-۴)$$

این نامساوی را می‌توان به صورت زیر بازنویسی کرد:

$$\Pr[\mathcal{M}(x) \in \mathcal{S}] - e^\alpha \Pr[\mathcal{M}(x') \in \mathcal{S}] \leqslant ۰ \qquad (۶-۴)$$

از آن‌جایی که این رابطه برای *تمام* $\mathcal{S}$ها برقرار است، سوپریمم آن نیز باید کوچکتر یا مساوی صفر باشد. اما طبق تعریف E_γ در معادله ۴-۱، این سوپریمم دقیقاً همان E_{e^α} است. چون E_γ نمی‌تواند منفی باشد (با انتخاب $\mathcal{S} = \emptyset$ مقدار حداقل صفر است)، پس حتماً برابر صفر است.

جهت دوم ($\Leftarrow$): فرض کنید $E_{e^\alpha}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) = ۰$. طبق تعریف سوپریمم، برای هر مجموعه دلخواه $\mathcal{S}$:

$$\Pr[\mathcal{M}(x) \in \mathcal{S}] - e^\alpha \Pr[\mathcal{M}(x') \in \mathcal{S}] \leqslant ۰ \qquad (۷-۴)$$

که بلافاصله نتیجه می‌دهد:

$$\frac{\Pr[\mathcal{M}(x) \in \mathcal{S}]}{\Pr[\mathcal{M}(x') \in \mathcal{S}]} \leqslant e^\alpha \qquad (۸-۴)$$

این دقیقاً همان تعریف α-LDP است. □

این قضیه ساده اما بنیادین، یک تفسیر هندسی دقیق از محرمانگی ارائه می‌دهد: **α-LDP یعنی توزیع‌های خروجی چنان به هم نزدیک باشند که هیچ بخشی از دامنه نتواند نسبت درست‌نمایی بیشتر از e^α ایجاد کند.**

۴-۴ بهبود کران‌های انقباض

در فصل ۳ دیدیم که دوچی [؟] کران زیر را برای انقباض KL ارائه کرد:

$$D_{KL}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) \leqslant ۴(e^\alpha - ۱)^۲ \tag{۴-۹}$$

حال با استفاده از چارچوب E_γ، می‌توانیم کران‌های بسیار دقیق‌تری استخراج کنیم. آسوده و همکاران [؟] نشان داده‌اند که اگر شرط $E_{e^\alpha} = ۰$ برقرار باشد، می‌توان کران‌های انقباض برای سایر fـ واگرایی‌ها را از طریق بهینه‌سازی محدب به دست آورد.

قضیه‌ی ۴-۲ (کران دقیق انقباض KL) *اگر $\mathcal{M}$ یک مکانیزم α-LDP باشد، آنگاه:*

$$D_{KL}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) \leqslant \frac{e^\alpha - ۱}{e^\alpha + ۱} \cdot (e^\alpha - ۱) \tag{۴-۱۰}$$

برای مقادیر کوچک α (رژیم محرمانگی بالا)، این کران به $\alpha^۲/۲$ میل می‌کند که ۴ برابر کوچکتر (بهتر) از کران دوچی است.

تحلیل مقایسه‌ای: بیایید رفتار دو کران را در $\alpha \to ۰$ بررسی کنیم:

- کران دوچی: $۴(e^\alpha - ۱)^۲ \approx ۴\alpha^۲$
- کران مبتنی بر E_γ: $\frac{\alpha^۲}{۲} = \frac{\alpha}{۲} \cdot \alpha \approx$ (چون $\frac{e^\alpha-۱}{e^\alpha+۱} \approx \tanh(\alpha/۲) \approx \alpha/۲$)

این بهبود ضریب ثابت (از ۴ به ۰/۵) در تحلیل‌های مینیماکس بسیار حیاتی است و نشان می‌دهد که «اندازه نمونه موثر» واقعی می‌تواند تا ۸ برابر بهتر از چیزی باشد که آنالیزهای قبلی نشان می‌دادند.

۴-۵ تعمیم به محرمانگی تقریبی ((α,δ)-LDP)

یکی دیگر از قدرت‌های چارچوب E_γ، توانایی آن در توصیف ساده‌ی محرمانگی تقریبی است. یادآوری می‌کنیم که (α,δ)-LDP شرط زیر را دارد:

$$\Pr[\mathcal{M}(x) \in \mathcal{S}] \leqslant e^\alpha \Pr[\mathcal{M}(x') \in \mathcal{S}] + \delta \tag{۴-۱۱}$$

با بازنویسی این رابطه داریم:

$$\sup_{\mathcal{S}}(\Pr[\mathcal{M}(x) \in \mathcal{S}] - e^\alpha \Pr[\mathcal{M}(x') \in \mathcal{S}]) \leqslant \delta \tag{۴-۱۲}$$

که دقیقاً معادل است با:

$$E_{e^\alpha}(\mathcal{M}(\cdot|x)||\mathcal{M}(\cdot|x')) \leqslant \delta \quad (۱۳-۴)$$

نتیجه‌ی ۴-۳ *محرمانگی تقریبی* (α,δ)-*LDP* **دقیقاً معادل محدود کردن مقدار** E_{e^α} **ـ واگرایی توزیع‌های** *خروجی به مقدار* δ *است. این نتیجه نشان می‌دهد که* E_γ *ـ واگرایی طبیعی‌ترین زبان برای صحبت درباره محرمانگی تفاضلی (چه خالص و چه تقریبی) است.*

۴-۶ کاربرد در تخمین توزیع گسسته

برای نشان دادن کاربرد عملی این نتایج، مسئله تخمین توزیع احتمال روی یک دامنه kـ تایی را در نظر بگیرید. با استفاده از تکنیک‌های انقباض E_γ، می‌توان نشان داد که نرخ مینیماکس برای این مسئله تحت شرط α-LDP برابر است با:

$$\mathfrak{M}_n \asymp \frac{k}{n(e^\alpha - ۱)^۲} \quad (۱۴-۴)$$

در حالی که استفاده از تکنیک‌های کلاسیک (دوچی)، جمله‌ای به صورت $\frac{k}{n\alpha^۲}$ را پیشنهاد می‌کرد. تفاوت این دو عبارت در رژیم α بزرگ (محرمانگی کم) آشکار می‌شود؛ جایی که $(e^\alpha - ۱)^۲$ به صورت نمایی رشد می‌کند و نشان می‌دهد که دقت می‌تواند بسیار سریع‌تر از پیش‌بینی‌های قبلی بهبود یابد.

۴-۷ انقباض قوی برای خانواده‌ی fـ واگرایی‌ها

تا اینجا دیدیم که شرط α-LDP معادل صفر شدن E_{e^α}ـ واگرایی است. یک پرسش طبیعی و بسیار مهم این است: آیا این شرط بر روی سایر معیارهای فاصله (مثل $\chi^۲$ یا هلینجر) نیز انقباض ایجاد می‌کند؟ پاسخ مثبت است. در مقاله‌ی اخیر آسوده و ژانگ [؟]، نشان داده شده است که مکانیزم‌های موضعی خاصیت «انقباض قوی» را برای طیف وسیعی از واگرایی‌ها به ارمغان می‌آورند.

۴-۷-۱ کران دقیق برای واگرایی کای‌ـ دو ($\chi^۲$)

یکی از مهم‌ترین نتایج این پژوهش، ارائه‌ی یک ضریب انقباض دقیق برای واگرایی $\chi^۲$ است. اهمیت این واگرایی در آن است که کار با آن در محاسبات واریانس و کران‌های مینیماکس بسیار ساده‌تر از KL است.

قضیه‌ی ۴-۴ (انقباض $\chi^۲$) *فرض کنید $\mathcal{M}$ یک مکانیزم α-LDP باشد. برای هر دو توزیع ورودی P و Q، واگرایی کای- دو بین توزیع‌های خروجی با رابطه زیر محدود می‌شود:*

$$\chi^۲(\mathcal{M}P||\mathcal{M}Q) \leqslant \eta_\alpha \cdot \chi^۲(P||Q) \tag{۴-۱۵}$$

که در آن η_α ضریب انقباض بهینه است و برابر است با:

$$\eta_\alpha = \left(\frac{e^\alpha - ۱}{e^\alpha + ۱}\right)^۲ \tag{۴-۱۶}$$

تحلیل مجانبی: برای مقادیر کوچک α (رژیم محرمانگی بالا)، داریم:

$$\eta_\alpha \approx \left(\frac{۱ + \alpha - ۱}{۱ + \alpha + ۱}\right)^۲ \approx \left(\frac{\alpha}{۲}\right)^۲ = \frac{\alpha^۲}{۴} \tag{۴-۱۷}$$

این نتیجه بسیار قابل توجه است. یادآوری می‌کنیم که کران‌های کلاسیک دوچی (فصل ۳) ضریبی از مرتبه $O(\alpha^۲)$ داشتند، اما ضریب ۱/۴ در اینجا نشان‌دهنده‌ی یک انقباض بسیار شدیدتر است. این ضریب دقیقاً با ضریب انقباض «پاسخ تصادفی دودویی» برای واریانس مطابقت دارد و نشان می‌دهد که این کران برای کل کلاس مکانیزم‌های α-LDP «تایت» (Tight) است.

۴-۷-۲ تعمیم به سایر واگرایی‌ها

نویسندگان در [؟] نشان داده‌اند که این ضریب انقباض η_α تنها مختص $\chi^۲$ نیست، بلکه برای خانواده‌ای از واگرایی‌ها که خاصیت «تحدب مشترک» دارند (شامل فاصله هلینجر مجذور $H^۲$ و واگرایی KL) نیز صادق است.

نتیجه‌ی ۴-۵ *برای هر مکانیزم α-LDP، کران‌های زیر برقرار هستند:*

$$D_{KL}(\mathcal{M}P||\mathcal{M}Q) \leqslant \eta_\alpha \cdot D_{KL}(P||Q) \tag{۴-۱۸}$$

$$H^۲(\mathcal{M}P, \mathcal{M}Q) \leqslant \eta_\alpha \cdot H^۲(P, Q) \tag{۴-۱۹}$$

این یکسان‌سازی ضرایب انقباض، تحلیل مکانیزم‌های پیچیده را بسیار ساده می‌کند؛ زیرا کافیست فقط ضریب η_α را محاسبه کنیم.

۸-۴ نامساوی ون‌تریز خصوصی (Private van Trees Inequality)

اکثر تحلیل‌های موجود در ادبیات α-LDP (مانند کارهای دوچی)، بر روی «ریسک مینیماکس» (بدترین حالت) تمرکز دارند. اما در بسیاری از کاربردهای مدرن، ما به تحلیل‌های بیزی (Bayesian) علاقه‌مندیم، جایی که پارامتر مجهول θ دارای یک توزیع پیشین $\pi(\theta)$ است.

نامساوی ون‌تریز (van Trees) ابزاری کلاسیک برای کران‌دار کردن خطای بیزی بر اساس «اطلاعات فیشر» است. آسوده و ژانگ [؟] نسخه‌ی خصوصی‌شده‌ی این نامساوی را ارائه کرده‌اند که ابزاری نوین در جعبه‌ابزار تحلیل محرمانگی محسوب می‌شود.

قضیه‌ی ۴-۶ (نامساوی ون‌تریز موضعی) *فرض کنید می‌خواهیم پارامتر θ را از روی مشاهدات Z^n که خروجی یک مکانیزم α-LDP هستند، تخمین بزنیم. اگر $\hat{\theta}$ هر تخمین‌گر دلخواهی باشد، آنگاه میانگین مربعات خطای بیزی[۲] دارای کران پایین زیر است:*

$$\mathbb{E}_{(\hat{\theta}-\theta)^2} \geqslant \frac{1}{\mathbb{E}_{\mathcal{I}(\theta)} + \mathcal{I}_{prior}(\pi)} \tag{۴-۲۰}$$

نکته‌ی کلیدی اینجاست که در نسخه خصوصی، اطلاعات فیشر مشاهدات ($\mathcal{I}(\theta)$) با ضریب انقباض تضعیف می‌شود:

$$\mathcal{I}_{priv}(\theta) \leqslant \eta_\alpha \cdot \mathcal{I}_{orig}(\theta) \tag{۴-۲۱}$$

که در آن $\mathcal{I}_{orig}$ اطلاعات فیشر داده‌های خام است.

تفسیر: این نامساوی به زبان ساده می‌گوید: «در دنیای α-LDP، هر بیت اطلاعات فیشر که از داده‌ها می‌گیرید، به اندازه‌ی $\eta_\alpha \approx \alpha^2/4$ تضعیف می‌شود.» این نتیجه، اثبات حدود پایین برای مسائل تخمین پارامتر را بسیار ساده می‌کند. به جای درگیر شدن با لم‌های پیچیده‌ی اسود یا فانو، کافیست اطلاعات فیشر مسئله‌ی اصلی را محاسبه کنیم و در ضریب η_α ضرب کنیم.

۹-۴ کاربردهای نوین و بهبود نرخ‌ها

استفاده از کران‌های انقباض قوی (بخش ۴-۷) و نامساوی ون‌تریز خصوصی (بخش ۴-۸) منجر به بهبود نتایج در مسائل کلاسیک می‌شود.

[2] Bayesian Mean Square Error

به عنوان مثال، در مسئله‌ی **تخمین چگالی غیرپارامتری** برای کلاس توزیع‌های هموار (کلاس هولدر با پارامتر β)، استفاده از این ابزارهای جدید نشان می‌دهد که نرخ خطای بهینه دقیقاً برابر است با:

$$R_{opt} \asymp \left(\frac{۱}{n\alpha^۲}\right)^{\frac{۲\beta}{۲\beta+۱}} \qquad (۲۲-۴)$$

اگرچه مرتبه‌ی کلی نرخ همگرایی مشابه نتایج دوچی است، اما ضرایب ثابت بهبود یافته‌اند و مهم‌تر از آن، اثبات با استفاده از انقباض $\chi^۲$ بسیار کوتاه‌تر و مستقیم‌تر از روش‌های مبتنی بر KL است. این امر نشان‌دهنده‌ی برتری رویکرد مبتنی بر E_γ و انقباض قوی در تحلیل سیستم‌های محرمانگی تفاضلی است.

فصل ۵

نتیجه‌گیری

واژه‌نامه

الف

ب

پ

پرس‌وجو .... Query

پایگاه‌داده .... Database

ت

ج

ح

خ

د

داده .... Data

دودویی .... Binary

ر

ز

س

ش

ص

غ

ف

ق

ک

گ

ل

م

مجموعه .... Set
متصدی مورد اعتماد .... Trusted Curator
مکانیزم تصادفی .... Randomized Mechanism
محرمانگی تفاضلی .... Differential Privacy
مطلقاً پیوسته .... Absolutely Continuous

ن

و

واگرایی .... Divergence

هـ

همسایه .... Ajacent
همسایگی .... Adjacency

ی

پیوست آ

مطالب تکمیلی

Abstract

We present a standard template for typesetting theses in Persian. The template is based on the XePersian package for the LaTeX typesetting system. This write-up shows a sample usage of this template.

Keywords: Thesis, Typesetting, Template, XePersian

Sharif University of Technology

Department of Mathematics

M.Sc. Thesis

Information-Theoretic Analysis of Local Differential Privacy and its Statistical Applications

By:

Firoozeh Abrishami

Supervisor:

Dr. Javad Ebrahimi Boroujeni

February 2026