

توسیع ایده آنتروپی ماکسیمم برای اندازه‌های اطلاع تعمیم‌یافته

منیژه صانعی طبس، غلامرضا محتشمی برزادران
گروه آمار، دانشگاه فردوسی مشهد

تاریخ دریافت: ۱۳۹۴/۹/۱۶ تاریخ آخرین بازنگری: ۱۳۹۶/۶/۱۵

چکیده: آنتروپی رنی ماکسیمم و آنتروپی تی‌سالیس ماکسیمم توسعه ایده آنتروپی ماکسیمم به رده بزرگتری از آنتروپی‌ها شانون است. در این مقاله ضمن معرفی آنتروپی رنی ماکسیمم به برخی از توزیع‌های خاص که آنتروپی رنی را ماکسیمم می‌کند، اشاره می‌شود. توزیع‌های دارای آنتروپی رنی ماکسیمم به شکل توزیع‌های توانی هستند. برخی از ویژگی‌های توزیع‌های توانی ارائه و نمایش جدیدی از آنتروپی رنی حاصل می‌شود. به بحث مینیمم اندازه اطلاع رنی اشاره و در این زمینه نیز نکاتی مطرح گردیده است. همچنین شکل دیگری از اندازه اطلاع رنی به دست می‌آید. در ادامه کاربری اقتصادی از آنتروپی رنی ماکسیمم بررسی و ضمن یادآوری اندازه اطلاع سیزار، فرم کلی توزیع‌های مینیمم‌کننده این اندازه اطلاع تعمیم‌یافته نتیجه‌گیری می‌شود.

واژه‌های کلیدی: آنتروپی ماکسیمم، آنتروپی رنی ماکسیمم، آنتروپی تی‌سالیس ماکسیمم، مینیمم اندازه اطلاع رنی.

۱ مقدمه

بهینه‌سازی همواره مورد توجه علوم زیادی بوده است. در نظریه اطلاع نیز این مبحث جایگاه خاص خود را دارد. یکی از روش‌های بهینه‌سازی اصل آنتروپی ماکسیمم به روش لاگرانژ است. سؤالی که مطرح می‌شود این است که چرا آنتروپی بهینه می‌شود؟

اصل آنتروپی ماکسیم جینز (۱۹۵۷)

بین همه توزیع‌های احتمال سازگار با یک سری از محدودیت‌های داده شده، توزیعی که دارای ماکسیمم عدم قطعیت است انتخاب می‌شود.

هدف، کاهش دادن عدم قطعیت است در حالی که سعی می‌شود ماکسیمم شود به نظر می‌رسد یک تناقض وجود دارد! عدم قطعیت را به کمک اطلاعاتی که در دسترس است کاهش داده و بنابراین ما رابه ماکسیمم کاهش در عدم قطعیت، به کمک همه اطلاعاتی که فراهم شده‌است، می‌رساند. این عدم قطعیت باقیمانده ناشی از اطلاعاتی است که فراهم نیست، یا هیچ اطلاعی در مورد آن داده نشده است. پس لازم است عدم قطعیت در مورد آنچه که نامعلوم است ماکسیمم شود که این درست‌ترین کار برای انجام است. چگونگی یافتن توزیعی که آنتروپی را ماکسیمم می‌کند با توجه به محدودیت‌هایی که علاوه بر تابع (احتمال) چگالی بودن لازم است توسط کاگان و همکاران (۱۹۷۳) و کاپور (۱۹۸۹) بیان گردید. توزیع دارای آنتروپی ماکسیمم معمولاً به صورت تابع نمایی از محدودیت‌ها بیان می‌شود یکی از این حالت‌ها زمانی است که گشتاورهای مراتب مختلف مشخص باشد که به ازای دامنه تغییرات متغیر تصادفی توزیع‌های مشهوری را به عنوان توزیع آنتروپی ماکسیمم نتیجه می‌دهد. منصوری و پاشا (۱۳۸۸) مباحثی در زمینه توزیع توام ماکسیمم آنتروپی مطرح نمودند. دستمرد و همکاران (۱۳۹۰) برخی از ویژگی‌های تعمیم‌هایی از توزیع‌های گسسته را به کمک آنتروپی ماکسیمم بررسی کردند. به منظور توسیع ایده آنتروپی ماکسیمم برای آنتروپی رنی و آنتروپی تی‌سالیس، کستا و همکاران (۲۰۰۲ و ۲۰۰۶)، باشکیرو (۲۰۰۴ و ۲۰۰۶)، جانسون و ویگنات (۲۰۰۷)، براودی و همکاران (۲۰۰۷)، و ناگی و رمرا (۲۰۰۹)، تعبیرها و مشخصه‌سازی‌هایی بر اساس آنتروپی رنی ماکسیمم و یا آنتروپی تی‌سالیس ماکسیمم در حالت‌های یک‌متغیره و چندمتغیره ارائه داده‌اند. اگر X یک متغیر تصادفی گسسته با تابع جرم احتمال $p_X(x)$ باشد، آنگاه انواع آنتروپی آن عبارتند از:

آنتروپی شانون (۱۹۴۸):

$$H(X) = - \sum_x p_X(x) \log p_X(x) \quad (1)$$

آنتروپی رنی (۱۹۶۱):

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \sum_x p_X^\alpha(x), \quad \alpha > 0, \alpha \neq 1 \quad (2)$$

آنتروپی تی‌سالیس (۱۹۸۸):

$$S_{\alpha}(X) = \frac{1}{\alpha - 1} \left[1 - \sum_x p_X(x)^{\alpha} \right], \quad \alpha > 0, \alpha \neq 1 \quad (3)$$

اطلاع کولبک-لیبلر (۱۹۵۱):

از میان دو مدل احتمال F و G یکی باید به‌عنوان مدلی برای متغیر تصادفی X برگزیده شود. برای مشاهده $X = x$ ، اطلاع کولبک-لیبلر موجود در هر مشاهده $X = x$ در پشتیبانی از F در مقابل G عبارت است از:

$$D(f||g) = \int_0^{\infty} \log \frac{f(x)}{g(x)} dF(x) \quad (4)$$

آنتروپی رنی نسبی یا اندازه اطلاع رنی:

۱۱ با ایده از کولبک (۱۹۵۹)، فرض کنید از میان دو مدل احتمال F و G یکی باید به‌عنوان مدلی برای متغیر تصادفی X انتخاب شود. اگر مشاهده $X = x$ را داشته باشیم آنگاه اطلاع رنی موجود در هر مشاهده $X = x$ در پشتیبانی از F در مقابل G عبارت است از:

$$D_{\alpha}(f||g) = \frac{1}{1 - \alpha} \log \int_0^{\infty} \left(\frac{f(x)}{g(x)} \right)^{\alpha - 1} dF(x), \quad (5)$$

به شرط آنکه F نسبت به G مطلقاً پیوسته باشد. رنی این کمیت را اطلاع رنی مرتبه α معرفی نمود. در این مقاله به مبحث آنتروپی رنی ماکسیم پرداخته و نکاتی در این زمینه اشاره می‌شود. فرم کلی توزیع دارای آنتروپی تی‌سالیس ماکسیم را به‌دست آورده و سپس موضوع مینیمم اندازه اطلاع رنی مورد بررسی قرار می‌گیرد. با تعریف اندازه اطلاع سیزار، شکل توزیعی که این اندازه اطلاع را تحت محدودیت‌های مشخص مینیمم می‌کند تعیین می‌شود. همچنین کاربری اقتصادی از آنتروپی رنی ماکسیم ارائه خواهد شد.

۲ آنتروپی رنی ماکسیم

در بسیاری از مواقع با در اختیار داشتن مقادیر گشتاورهای مراتب مختلف یک توزیع، یافتن توزیع با آنتروپی ماکسیم، مورد نظر است.

۱۰۴ آنتروپی ماکسیم برای اندازه‌های اطلاع تعمیم یافته

قضیه ۱ (کاگان وهمکاران، ۱۹۷۳): متغیر تصادفی X با تابع چگالی $f(x) > 0, \forall x \in D$ که $\int_D f(x) dx = 1$ را در نظر بگیرید. اگر توابع $h_i(x), i = 1, \dots, k$ طوری تعریف شده باشند که $E(h_i(X)) = \theta_i$ و θ_i ها ثابت باشند، آنگاه توزیع دارای آنتروپی ماکسیم به صورت

$$f(x) = e^{c_0 + c_1 h_1(x) + \dots + c_k h_k(x)}$$

است که در آن c_0, c_1, \dots, c_k به کمک محدودیت‌های ذکر شده به دست می‌آیند.

۲،۱ آنتروپی رنی ماکسیم تحت محدودیت کلی $E(g(X)) = u$

فرض کنید $f(x)$ تابع چگالی احتمال نامعلوم بر اعداد حقیقی مثبت باشد. تابع $g(x)$ با محدودیت کلی $E(g(X)) = u$ را در نظر بگیرید. تابع چگالی احتمال $f(x)$ که تحت این محدودیت آنتروپی رنی رابطه (۲) را ماکسیم می‌کند، باید تعیین شود. با تشکیل معادله لاگرانژ اگر فرض شود $I = \int_a^b F(x, f(x), f'(x)) dx$ که در آن F یک تابع معلوم است. به کمک معادله لاگرانژ اوایلر، تابع $f(x)$ که I را ماکسیم کند در رابطه $\frac{\partial F}{\partial f(x)} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial f'(x)} \right) = 0$ صدق می‌کند. در این مسئله، تابع $f'(x)$ مجهول است، تابع زیر انتگرال، تابعی از x و $f(x)$ است و فقط اولین عبارت این رابطه مشخص است. برای اطلاع بیشتر در این زمینه به (وحیدیان کامیاد و بزرگ‌نیا ۱۳۷۲ ص ۲۴۷) مراجعه شود. حال با مشتق‌گیری از طرفین معادله لاگرانژ، تابع $f(x)$ به صورت

$$f(x) = \frac{1}{Z_\beta} \left[1 - \frac{1-\alpha}{\alpha} \beta(u - g(x)) \right]^{\frac{1}{\alpha-1}} \quad (6)$$

به دست می‌آید که در آن $Z_\beta = \left[\int_0^\infty f^\alpha(x) dx \right]^{\frac{1}{1-\alpha}}$. در حالتی که چندین محدودیت به صورت $Eg_i(X) = u_i, i = 1, \dots, n$ داده شده باشند، تابع $f(x)$ که آنتروپی رنی را ماکسیم می‌کند، به صورت

$$f(x) = \frac{1}{B(\alpha, u_1, \dots, u_n)} \left[1 - \frac{1-\alpha}{\alpha} \sum_{i=1}^n \beta_i(u_i - g_i(x)) \right]^{\frac{1}{\alpha-1}} \quad (7)$$

است که در آن

$$B(\alpha, u_1, \dots, u_n) = \int_0^\infty \left[1 - \frac{1-\alpha}{\alpha} \sum_{i=1}^n \beta_i (u_i - g_i(x)) \right]^{\frac{1}{\alpha-1}} dx$$

یکی از حالاتی که توابع $g_i(x)$ می‌توانند اختیار کنند، توابعی به صورت $g_i(x) = x^{k_i}$ است و اگر $k_i = i$ محدودیت‌ها گشتاوری هستند که در ادامه مورد بررسی قرار می‌گیرند. فرض کنید تابع چگالی احتمال نامعلوم روی اعداد حقیقی مثبت تعریف شده است. همچنین k امین گشتاور توزیع $f(x)$ معلوم است. آنچه باید تعیین شود تابع چگالی احتمال $f(x) = \left[\int_0^\infty f(x)^\alpha dx \right]^{\frac{1}{\alpha-1}} \left[1 - \frac{1-\alpha}{\alpha} \beta (u - x^k) \right]^{\frac{1}{\alpha-1}}$ است که با معلوم بودن گشتاور k ام آنتروپی رنی را ماکسیم می‌کند. در حالت کلی می‌توان وضعیتی را در نظر گرفت که چندین محدودیت گشتاوری، داده شده باشد:

در این مورد نمایش (۷) به صورت

$$f(x) = \left[\int_0^\infty f(x)^\alpha dx \right]^{\frac{1}{\alpha-1}} \left[1 - \frac{1-\alpha}{\alpha} \sum_{i=1}^n \beta_i (u_i - x^{k_i}) \right]^{\frac{1}{\alpha-1}} \quad (۸)$$

نوشته می‌شود. برخلاف حالتی که فقط یک محدودیت داده شده، انتگرال‌گیری از طرف راست رابطه (۸) به سادگی انجام نمی‌گیرد (به جز برای حالاتی که $k_1 = 1$ ، $k_2 = 2$ یعنی انتگرال با توابع فوق هندسی داده شده باشد). بنابراین در حالت کلی باید به استفاده از روش‌های عددی برای یافتن ضرایب لاگرانژ، روی آورد. در اینجا ماکسیم آنتروپی رنی برای توزیع‌های n متغیره مورد بررسی قرار می‌گیرند.

تعریف ۱ (کستا و همکاران، ۲۰۰۶): برای $\alpha > \frac{n}{n+1}$ و $\alpha \neq 1$ چگالی احتمال بردار تصادفی $X = (X_1, \dots, X_n)$ با ماتریس کواریانس C به صورت

$$g_{\alpha, c} = A_\alpha (1 - (\alpha - 1) \beta x^T C^{-1} x)_+^{\frac{1}{\alpha-1}}, \quad x = (x_1, \dots, x_n) \quad (۹)$$

در نظر بگیرید که در آن $\beta = \frac{1}{2\alpha - n(1-\alpha)}$ ، $x_+ = \max(0, x)$ و

$$A_\alpha = \begin{cases} \frac{\gamma(\frac{1}{1-\alpha})(\beta(1-\alpha))^{\frac{n}{\alpha-1}}}{\gamma(\frac{1}{1-\alpha})\pi^{\frac{n}{\alpha-1}}|C|^{\frac{1}{\alpha-1}}} \frac{n}{n+1} & \frac{n}{n+1} < \alpha < 1 \\ \frac{\gamma(\frac{\alpha}{\alpha-1} + \frac{n}{\alpha-1})(\beta(\alpha-1))^{\frac{n}{\alpha-1}}}{\gamma(\frac{\alpha}{\alpha-1})\pi^{\frac{n}{\alpha-1}}|C|^{\frac{1}{\alpha-1}}} & \alpha > 1 \end{cases} \quad (۱۰)$$

۱۰۶ آنتروپی ماکسیم برای اندازه‌های اطلاع تعمیم‌یافته

به‌ازای $1 < \alpha < \frac{n}{n+2}$ تابع چگالی $g_{\alpha,c}$ چگالی n متغیره t استیودنت با درجه آزادی n $m = \frac{2}{1-\alpha} - n$ و ماتریس کواریانس C است و با $T(m, n, C)$ نشان داده می‌شود. به‌ازای $\alpha > 1$ نیز تابع چگالی احتمال n متغیره r استیودنت با درجه آزادی n $p = \frac{2\alpha}{\alpha-1} + n$ و ماتریس کواریانس C به‌دست می‌آید که با $R(p, n, C)$ نشان می‌دهند.

توجه شود که:

$$\lim_{\alpha \rightarrow 1} g_{\alpha C}(x) = g_{1C}(x) = ((2\pi)^n |C|)^{-\frac{1}{2}} \exp(-x^T C^{-1} x / 2).$$

به عبارت دیگر وقتی α به سمت یک میل می‌کند، $g_{\alpha,c}(x)$ به تابع چگالی گاوسی میل می‌کند.

قضیه ۲ (جانسون و ویگنات، ۲۰۰۷): برای $\alpha > \frac{n}{n+2}$ و ماتریس کواریانس C (مثبت متقارن و منتهای)، از بین همه توابع چگالی احتمال $f(x)$ که دارای میانگین صفر و در محدودیت

$$\int_{\Omega_{\alpha,c}} x x^T f(x) dx = C$$

صدق می‌کنند، تابع $g_{\alpha,c}(x)$ به‌طور یکتا آنتروپی رنی را ماکسیم می‌کند.

در حالت یک‌متغیره تابع چگالی احتمال $g_{\alpha,c}(x)$ که آنتروپی رنی را تحت محدودیت $E(X^2) = u$ ماکسیم می‌کند با قراردادن $n = 1$ در تابع چگالی احتمال (۱۰) برای $1 < \alpha < \frac{1}{2}$ به‌دست می‌آید.

۲,۲ مشخصه‌هایی از آنتروپی رنی ماکسیم

اصل آنتروپی ماکسیم جینز پیشنهاد می‌کند که توزیع احتمال با حداقل اربیی، توزیعی با آنتروپی ماکسیم درمقایسه با همه اطلاعات پیشین است. فرض کنید دو تابع چگالی احتمال f و g داشته باشیم و هدف محاسبه میزان رابطه بین دو توزیع باشد پیشنهاد می‌شود با پارامتری کردن $D(f||g)$ آنتروپی تابع چگالی f ماکسیم شود. پس معادله به‌صورت

$$\begin{cases} \max_f - \int_0^\infty f(x) \log f(x) dx \\ \text{s.t} \\ D(f||g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx = \theta \end{cases} \quad (11)$$

است و تعریف $D(f||g)$ مستلزم این است که f نسبت به g مطلقاً پیوسته باشد. از طرفی $f(x)$ یک تابع چگالی احتمال است. تابع

$$f(x) = \frac{g(x)^\alpha}{\int_0^\infty g(x)^\alpha dx} \quad (12)$$

دارای آنتروپی رنی ماکسیمم است. این فرم تابع $f(x)$ را تابع چگالی همراه می‌نامند. بنابراین این نتیجه به‌دست آمد که مسئله ماکسیمم آنتروپی (۱۱) برای مقدار بهینه آنتروپی رنی رابطه (۱۲)، یک تابع خطی از محدودیت و آنتروپی رنی است. حال مسئله ابتدایی را با یک محدودیت اضافی در نظر بگیرید یعنی علاوه بر محدودیت اطلاع کولبک-لیبلر، محدودیت میانگین نیز منظور شود. همان‌طور که در بالا اشاره شد، آنتروپی بهینه تحت اولین محدودیت تابعی خطی از آنتروپی رنی است. بنابراین تحت محدودیت اضافی کفایت آنتروپی رنی تحت رابطه (۱۲) ماکسیمم شود. در این صورت:

$$f(x) = \left[\int_0^\infty g(x)^\alpha dx \right]^{\frac{1}{\alpha-1}} [1 + \beta(\alpha-1)(x-m)]^{-\frac{1}{\alpha-1}}$$

این تابع چگالی توزیع پاراتو تعمیم‌یافته با پارامتر شکل $(\alpha-1)$ و پارامتر مقیاس β است. توزیع دارای ماکسیمم آنتروپی رنی تحت محدودیت میانگین، از خانواده توزیع پاراتو تعمیم‌یافته است (صانعی وهمکاران، ۱۳۸۹). ولی پارامتر شکل آن برابر $\frac{1}{\alpha} - 1$ است. این تفاوت ناشی از این است که در آن مسئله، برای تعریف محدودیت میانگین از تابع چگالی معمولی استفاده شد ولی در اینجا از تابع چگالی همراه استفاده شده‌است. اگر حد تابع $f(x) \propto [1 + \beta(\alpha-1)(x-m)]^{-\frac{\alpha}{\alpha-1}}$ را وقتی $\alpha \rightarrow 1$ در نظر گرفته شود توزیع $f(x)$ به شکل توزیع نمایی است که ماکسیمم آنتروپی شانون تحت محدودیت میانگین روی تکیه‌گاه $(0, \infty)$ نیز همان توزیع نمایی را نتیجه می‌دهد. اگر به جای آنتروپی ماکسیمم تحت محدودیت اطلاع کولبک-لیبلر، مسئله به صورت

$$\begin{cases} \min_f D(f||g), \\ s.t \\ H(P) = \theta' \end{cases} \quad (13)$$

۱۰۸ آنتروپی ماکسیمم برای اندازه‌های اطلاع تعمیم یافته

باشد، با فرض $\lambda = \frac{1}{\beta}$ داریم: $f(x) = g(x)^{\frac{1}{\lambda-1}} e^{\frac{\lambda-1}{\lambda} x} = \frac{1}{Z(\lambda)^{\frac{1}{\lambda-1}}} g(x)^{\frac{\lambda}{\lambda-1}}$. دو مسئله (۱۱) و (۱۳) منجر به تابع چگالی $f(x)$ یکسان می‌گردد (برچر، ۲۰۰۸). حال این موضوع، برای آنتروپی رنی و اندازه اطلاع رنی (رابطه (۵)) مورد بررسی قرار می‌گیرد.

$$\begin{cases} \max_f H_\alpha(f). \\ s.t \\ D_\alpha(f||g) = \theta \end{cases}$$

با مشتق‌گیری از معادله لاگرانژ نسبت به $f(x)$ داریم:

$$\frac{f(x)^{\alpha-1}}{\int_0^\infty f(x)^\alpha dx} \left[\frac{\int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx + \beta g(x)^{1-\alpha} \int_0^\infty f(x)^\alpha dx}{\int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx} \right] = 1 + \beta.$$

در ادامه مسئله مینیمم اطلاع رنی به شرط معلوم بودن آنتروپی رنی بررسی می‌شود:

$$\begin{cases} \min D_\alpha(f||g). \\ s.t \\ H_\alpha(f) = \theta' \end{cases}$$

معادله لاگرانژ را تشکیل داده، با مشتق‌گیری نسبت به $f(x)$ و با قرار دادن $\beta' = \frac{1}{\beta}$ روابط

$$\frac{f(x)^{\alpha-1}}{\int_0^\infty f(x)^\alpha dx} \left(\frac{\frac{1}{\beta'} g(x)^{1-\alpha} \int_0^\infty f(x)^\alpha dx + \int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx}{\int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx} \right) = \frac{1 + \beta}{\beta}$$

$$\frac{f(x)^{\alpha-1}}{\int_0^\infty f(x)^\alpha dx} \left(\frac{\beta' g(x)^{1-\alpha} \int_0^\infty f(x)^\alpha dx + \int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx}{\int_0^\infty f(x)^\alpha g(x)^{1-\alpha} dx} \right) = (1 + \beta')$$

به دست می‌آیند. با انجام محاسبات لازم ملاحظه می‌شود که همانند آنتروپی شانون برای آنتروپی رنی نیز حل هر یک از دو معادله ماکسیمم آنتروپی رنی تحت محدودیت اندازه اطلاع رنی و نیز مینیمم اطلاع رنی تحت محدودیت آنتروپی رنی به جواب یکسانی منجر می‌شود.

۲,۳ خواص توزیع‌های توانی حاصل از آنتروپی رنی ماکسیم

با توجه به اینکه توزیع‌های حاصل از آنتروپی رنی ماکسیم، به صورت توزیع‌های توانی هستند (بشپرو، ۲۰۰۴) در این جا بعضی از این توزیع‌ها تحت محدودیت کلی $E[g(X)] = u$ بررسی می‌شود. برای یادآوری با در نظر گرفتن توزیع (۶) و $f(x)$ باید در شرط تابع چگالی بودن صدق کند باید $f(x)$ تساوی را به گزاره ای درست تبدیل نماید. به عبارت دیگر اگر توان عبارت داخل انتگرال به اندازه یک واحد انتقال پیدا کند، Z_β تغییر نمی‌کند. اگر چه در حالتی که $\alpha \rightarrow 1$ هر دو صورت زیر انتگرال تابع $e^{-\beta(g(x)-u)}$ را نتیجه می‌دهند. توجه داشته باشید دامنه α بسته به تابع $g(x)$ به طوری که انتگرال‌های دو طرف تساوی بالا تعریف شوند، تعیین می‌شود. اکنون رابطه $Z_\beta = (\int_0^\infty f^\alpha(x) dx)^{\frac{1}{1-\alpha}}$ را در نظر بگیرید. اگر از طرفین آن لگاریتم گرفته و با تعریف آنتروپی رنی مقایسه کنید نتیجه همان آنتروپی رنی است.

$$\log Z_\beta = \frac{1}{1-\alpha} \log \int_0^\infty f^\alpha(x) dx = H_\alpha(f(x))$$

به عبارت دیگر لگاریتم عامل نرمال‌کننده Z_β خود، آنتروپی رنی را به دست می‌دهد. حال اگر تابع (۶) در تعریف آنتروپی رنی جایگزین شود با استفاده از رابطه (۱۲) عبارت

$$H_\alpha(f(x)) = \log \int_0^\infty \left(1 - \frac{1-\alpha}{\alpha} \beta(u - g(x))\right)^{\frac{\alpha}{\alpha-1}} dx \quad (14)$$

حاصل می‌شود که صورت دیگری از آنتروپی رنی را نمایش می‌دهد.

آنتروپی تی‌سالیس ماکسیم

آنتروپی تی‌سالیس (۴) صورت تعمیم‌یافته دیگری از آنتروپی شانون است که به ازای $\alpha = 1$ آن را نتیجه می‌دهد. توزیع با ماکسیم آنتروپی تی‌سالیس تحت محدودیت‌های تابع چگالی بودن و $E(g(X)) = u$ همانند آنتروپی رنی، با روش لاگرانژ به صورت

$$f(x) = \left[\frac{1-\alpha}{\alpha} \left(\lambda + \sum_{i=1}^n \beta_i g_i(x) \right) \right]^{\frac{1}{\alpha-1}}$$

حاصل می‌شود که ضرایب λ و $\beta_i, i = 1, \dots, n$ تحت محدودیت‌های ذکر شده محاسبه می‌شوند.

۳ مینیمم اندازه اطلاع رنی

اصل مینیمم فاصله یا اصل مینیمم اطلاع، نیز یک اصل بهینه‌سازی است. رابطه‌ای نزدیک و تنگاتنگ بین مینیمم اندازه‌های اطلاع و اصل آنتروپی ماکسیم ثابت می‌شود. تاکید اصل مینیمم اطلاع روی فاصله یک تابع چگالی احتمال f از تابع چگالی احتمالی دیگری مانند g است. یک اندازه که برای ماهیت خاصی دارد، فاصله توزیع داده شده از توزیع یکنواخت است. اهمیت این اندازه از این جا ناشی می‌شود که این توزیع دارای بیشترین عدم قطعیت است. پس هر توزیع که به توزیع یکنواخت نزدیک تر باشد دارای عدم قطعیت بیشتری خواهد بود. بنابراین مینیمم کردن این اندازه تحت یک محدودیت معادل با ماکسیم کردن آنتروپی متناظر تحت همان محدودیت خواهد بود.

برودی و همکاران (۲۰۰۷) با استفاده از توزیع آنتروپی رنی ماکسیم تابع چگالی بدون مخاطره برای قیمت نهایی دارایی تحت تملک به دست آوردند. در راستای کار آنها در این تحقیق بر اساس مینیمم اندازه اطلاع رنی فرمول کلی تری برای تابع چگالی یاد شده یافته که کاربردهای مناسبی از آن در اقتصاد و بیمه است. در ساده ترین حالت اطلاع رنی، تحت محدودیت‌های گشتاوری باید مینیمم شود. تابع چگالی احتمال نامعلوم $f(x)$ تعریف شده بر اعداد حقیقی مثبت که با فرض معلوم بودن n گشتاور

$$\int_0^{\infty} x^{k_i} f(x) dx = u_i, \quad i = 1, \dots, n$$

اندازه اطلاع رنی را مینیمم می‌کند به صورت

$$f(x) = \frac{1}{Z_{\beta}} \left[1 + \frac{\alpha - 1}{\alpha} \sum_{i=1}^n \beta_i (x^{k_i} - u_i) \right]^{\frac{1}{\alpha-1}} g(x)$$

است و $n = 1$ معرف حالت یک متغیره خواهد بود.

۳,۱ رابطه بین مینیمم اندازه اطلاع رنی و ماکسیم آنتروپی رنی

فرض کنید تابع چگالی احتمال پیشین $g(x)$ داده شده است متناظر با اصل مینیمم اطلاع رنی تابع چگالی احتمال $f(x)$ را باید یافت که در محدودیت‌ها صدق کرده و به تابع چگالی احتمال $g(x)$ نزدیک ترین باشد. به عبارت دیگر $D_{\alpha}(f||g)$ تحت محدودیت‌های تابع چگالی بودن و $\int f(x)u(x)dx = a_r, r = 1, \dots, m$ مینیمم شود. اگر تابع چگالی احتمال $g(x)$ مشخص نباشد به جای آن توزیعی که دارای آنتروپی

رنی ماکسیمم است را در نظر گرفته و اگر هیچ محدودیتی داده نشده باشد، توزیع یکنواخت را به عنوان یک توزیع پیشین در نظر گرفته و $D_\alpha(f||u)$ مینیمم شود. که مینیمم کردن آن معادل با ماکسیمم کردن آنتروپی رنی است. بنابراین از این نگاه، آنتروپی رنی ماکسیمم به عنوان یک حالت خاصی از مینیمم اطلاع رنی در نظر گرفته می شود البته باید به تفاوت این دو مفهوم توجه داشت. در آنتروپی رنی ماکسیمم عدم قطعیت در مورد آنچه نمی دانیم با استفاده از همه اطلاعاتی که در دسترس است ماکسیمم می شود یعنی کمیت اساسی عدم قطعیت است. حال آنکه مینیمم اطلاع رنی مفهوم عدم قطعیت را در بر نمی گیرد بلکه نکته مهم فاصله احتمالی یک توزیع از توزیع دیگر است. مسئله یافتن یک توزیع f است که بین همه توزیع هایی که در محدودیت داده شده صدق می کنند نزدیکترین توزیع به g باشد. پس در آنتروپی ماکسیمم، عدم قطعیت ماکسیمم می شود در حالی که در مینیمم اطلاع رنی فاصله احتمالی یک توزیع از توزیعی معین، مینیمم می گردد.

صورت دیگری از اندازه اطلاع رنی اگر از طرفین رابطه $Z_\beta = (\int_0^\infty \frac{f(x)^\alpha}{g(x)^{\alpha-1}} dx)^{\frac{1}{1-\alpha}}$ لگاریتم گرفته و با تعریف اندازه اطلاع رنی مقایسه شود خواهیم داشت:

$$\ln(Z_\beta) = -D_\alpha(f||g)$$

از طرف دیگر

$$\begin{aligned} D_\alpha(f||g) &= \frac{1}{\alpha-1} \ln \left[\int_0^\infty \left(1 + \frac{\alpha-1}{\alpha} \beta(W(x)-u) \right)^{\frac{\alpha}{\alpha-1}} g(x)^\alpha dx - \alpha \ln Z_\beta \right] \\ &= \frac{1}{\alpha-1} \ln \left[\int_0^\infty \left(1 + \frac{\alpha-1}{\alpha} \beta(W(x)-u) \right)^{\frac{\alpha}{\alpha-1}} g(x)^\alpha dx + \alpha D_\alpha(f||g) \right] \\ \implies D_\alpha(f||g) &= \log \left[\int_0^\infty \left(1 + \frac{\alpha-1}{\alpha} \beta(W(x)-u) \right)^{\frac{\alpha}{\alpha-1}} g(x)^\alpha dx \right] \end{aligned}$$

که صورت دیگری از اندازه اطلاع رنی را وقتی محدودیت $E[W(X)] = u$ باشد، نشان می دهد.

۳,۲ خواص توزیع‌های توانی

یک محدودیت کلی‌تر نسبت به محدودیت قبلی (یعنی امید ریاضی یک تابع خطی $W(X)$) فرض کنید برابر u باشد. یعنی $\int_0^{\infty} W(x)f(x)dx = u$ پس توزیع مینیمم اطلاع رنی صورت

$$f(x) = \frac{1}{Z_{\beta}} \left[1 + \frac{\alpha - 1}{\alpha} \beta(W(x) - u) \right]^{\frac{1}{\alpha-1}} g(x) \quad (15)$$

را اختیار می‌کند که در آن $Z_{\beta} = \left(\int_0^{\infty} \frac{f(x)^{\alpha}}{g(x)^{\alpha-1}} dx \right)^{\frac{1}{1-\alpha}}$ عامل نرمال‌کننده است، به طوری که رابطه (۱۵) یک تابع چگالی احتمال باشد. همچنین $\frac{[f(x)Z_{\beta}]^{\alpha}}{Z_{\beta}g(x)^{\alpha-1}}$ نیز یک تابع چگالی است که گاهی اوقات به آن تابع چگالی همراه گفته می‌شود.

۴ اندازه اطلاع سیزار

فرض کنید $\phi : R^+ \rightarrow R^+$ یک تابع محدب باشد، سیزار در سال ۱۹۶۷ اندازه واگرایی ϕ تعمیم یافته، را به صورت $D_{\phi}(P||Q) = \sum_{i=1}^n q_i \phi\left(\frac{p_i}{q_i}\right)$ در فضای توزیع‌های احتمال P^n معرفی نمود. برحسب اینکه تابع ϕ به چه صورت تعریف شده باشد این اندازه هر یک از اندازه‌های واگرایی را در حالت خاص نتیجه می‌دهد. برای حالت پیوسته نیز به طور مشابه تعریف می‌شود.

۴,۱ مینیمم اندازه اطلاع تعمیم یافته

$f(x)$ تابع چگالی احتمال نامعلوم تعریف شده بر اعداد حقیقی مثبت است. هدف مینیمم اندازه واگرایی تعمیم یافته تحت معلوم بودن برخی از محدودیت‌ها است.

$$\begin{cases} \min_f (D_{\phi}(f(x)||g(x))) = \min_f \int_0^{\infty} g(x) \phi\left(\frac{f(x)}{g(x)}\right) dx \\ \text{s.t} \\ \int_0^{\infty} h_j(x) f(x) dx = u_j, j = 1, \dots, k \end{cases}$$

جدول ۰۱. برخی از اندازه‌های واگرایی و توابعی با اندازه‌های مینیم

$f(x)$	$\phi(t)$	اندازه
$g(x)e^{\lambda + \sum_{j=1}^k \beta_j h_j(x)}$	$t \log t$	کولیک- لیبیلر
$g(x) \left[\left(\frac{\alpha-1}{\alpha} \right) \left(\lambda + \sum_{j=1}^k \beta_j h_j(x) + \frac{1}{\alpha-1} \right) \right]^{\frac{1}{\alpha-1}}$	$\frac{t^\alpha - t}{\alpha-1}$	تی‌سالیس
$\frac{g(x)}{(1-2\lambda-2\sum_{j=1}^k \beta_j h_j(x))^r}$	$\frac{1}{r} (1-\sqrt{t})^2$	هلینجر
$g(x) \left[\left(\frac{\alpha-1}{\alpha} \right) \left(\lambda + \sum_{j=1}^k \beta_j h_j(x) \right) \right]^{\frac{1}{\alpha-1}}$	$\frac{r}{1-\alpha^r} [1 - t^{\frac{1+\alpha}{r}}]$	آلفا
$g(x) \left[\frac{\lambda}{r} + \frac{\sum_{j=1}^k \beta_j h_j(x)}{r} + 1 \right]$	$(t-1)^2$	خی دو
$\frac{1}{r} g(x) \left[\frac{1}{(\lambda + \sum_{j=1}^k \beta_j h_j(x))^r} \right]$	\sqrt{t}	باتاچاریا
$g(x) \left[\left(\frac{-r}{\lambda + \sum_{j=1}^k \beta_j h_j(x)} \right)^{\frac{1}{r}} - 1 \right]$	$\frac{r}{(t+1)^r}$	مثلی
$g(x) \left[\sqrt{\frac{r}{\lambda + \sum_{j=1}^k \beta_j h_j(x)}} - 1 \right]$	$\frac{rt}{t+1}$	هارمونیک

با تشکیل معادله لاگرانژ و حل آن رابطه

$$\phi' \left(\frac{f(x)}{g(x)} \right) - \lambda - \sum_{j=1}^k \beta_j h_j(x) = 0 \quad (16)$$

به دست می‌آید و بر حسب اینکه تابع ϕ چگونه تعریف شده باشد توابع مختلف $f(x)$ از رابطه (۱۶) حاصل می‌شود. (در حالت گسسته نیز به‌طور معادل تابع احتمال داریم وانتگرال به سیگما تبدیل می‌شود) برخی از این حالات در جدول ۱ خلاصه شده‌اند.

۵ کاربری اقتصادی آنتروپی رنی ماکسیم

آنتروپی رنی تاکنون مورد علاقه بسیاری از دانشمندان علوم مختلف بوده‌است که از جمله آن می‌توان به نظریه کدگذاری و ارتباطات، استخراج داده، بازیابی و طبقه‌بندی داده، آزمون فرضیه، ثبت و تماشای تصویر و ... اشاره کرد. به بحث کاربردی از آنتروپی رنی ماکسیم در اقتصاد پرداخته و براساس فرمولی که برودی و همکاران (۲۰۰۷) برای تابع چگالی بدون مخاطره قیمت‌ها به کمک ماکسیم آنتروپی رنی به دست آوردند. به مسئله تنظیم قیمت با استفاده از روش‌های ارائه شده پرداخته می‌شود برای سادگی از زیرنویس α در تابع $f_\alpha(x)$ صرف نظر شده و فرض کنید تابع $f(x)$ نشان دهنده تابع چگالی نامعلوم قیمت نهایی دارایی تحت

تملك باشد که باید برآورد شود. با استفاده از روش ماکسیم آنتروپی رنی تحت محدودیت‌های مشخص، تابع $f(x)$ به دست می‌آید.

تنظیم قیمت با استفاده از ماکسیم آنتروپی رنی

فرض کنید $f(x)$ تابع چگالی احتمال بدون مخاطره برای قیمت نهایی دارایی تحت تملك است. همچنین مقدار متوسط قیمت نهایی دارایی تحت تملك برابر با ضریبی از قیمت ابتدایی است. فرض کنید s قیمت ابتدایی دارایی تحت تملك باشد و t زمان پرداخت یا زمان انقضاء قرار داد باشد و همچنین r در اینجا تقریباً ثابت فرض می‌شود. بنابراین دومین محدودیت داده شده عبارت است از:

$$\int_0^{\infty} x f(x) dx = s_0 e^{rt} \quad (17)$$

علاوه بر این محدودیت‌ها، یک سری از قیمت‌های متناظر با یک سری از توابع پرداخت وجود دارند. که برای شروع کار ابتدا ساده ترین حالت را در نظر گرفته که فقط یک تابع پرداخت به شکل c_0 داریم. در نتیجه سومین محدودیت

$$\int_0^{\infty} (x - k_0)^+ f(x) dx = c_0 e^{rt} \quad (18)$$

است که در آن همان افت شکست قیمت قرارداد و $x^+ = \max\{0, x\}$ است. تحت این سه محدودیت برای ماکسیم کردن آنتروپی رنی اقدام می‌شود. فرض کنید λ و β_0 و β_1 به ترتیب نشان دهنده ضرایب لاگرانژ مربوط به محدودیت های تابع چگالی بودن و (۱۷) و (۱۸) باشند. در بخش دو ملاحظه شد که جواب آنتروپی رنی ماکسیم تحت محدودیت‌های یاد شده صورت

$$f(x) = [\tilde{\lambda} + \tilde{\beta}_0 x + \tilde{\beta}_1 (x - k_0)^+]^{\frac{1}{\alpha-1}} \quad (19)$$

را دارد.

$$\bar{c}_0 = c_0 e^{rt}, \quad \bar{s}_0 = s_0 e^{rt}, \quad \tilde{\lambda} = \frac{\lambda}{Z^{\alpha-1}}, \quad \beta_i = \frac{\beta_i}{Z^{\alpha-1}}, \quad i = 0, 1$$

با جایگذاری این ضرایب در معادله (۱۹) و انجام محاسبات نسبتاً طولانی تابع چگالی بدون مخاطره برای

قیمت دارایی برآورد شده است. بدون اینکه هیچ اطلاعی از توزیع دارایی در دسترس باشد. بعد از یافتن توزیع $f(x)$ به کمک آن می‌توان فرمول قیمت را برای سایر مشتقات متناظر با افت شکست k نیز برآورد نمود. که بر حسب اینکه $k \geq k_0$ یا $k \leq k_0$ در ادامه به آن پرداخته می‌شود. برای حالت $k \leq k_0$ برآورد قیمت تابع پرداخت متناظر با افت شکست k عبارت است از:

$$\bar{c}(k) = \frac{(\alpha - 1)^2}{\alpha(2\alpha - 1)} \left(\frac{1}{\tilde{\beta}_0 + \tilde{\beta}_1} \right)^2 [\tilde{\lambda} + \tilde{\beta}_0 k + \tilde{\beta}_1 (k - k_0)]^{\frac{2\alpha-1}{\alpha-1}}$$

و برای $k \geq k_0$ به روش مشابه بالا مقدار $\bar{c}(k)$ تعیین می‌شود. که بعد از محاسبات جبری رابطه

$$\begin{aligned} \bar{c}(k) = & \bar{s}_0 - \left(\frac{\alpha - 1}{\alpha} \right) \frac{\tilde{\beta}_1}{\tilde{\beta}_0 (\tilde{\beta}_0 + \tilde{\beta}_1)} (\tilde{\lambda} + \tilde{\beta}_0 k_0)^{\frac{\alpha}{\alpha-1}} k \\ & + \frac{(\alpha - 1)^2}{\alpha(2\alpha - 1)} \frac{1}{\tilde{\beta}_0^2} [(\tilde{\lambda} + \tilde{\beta}_0 k)^{\frac{2\alpha-1}{\alpha-1}} - \tilde{\lambda}^{\frac{2\alpha-1}{\alpha-1}}] \end{aligned}$$

نتیجه می‌شود. تابع $\bar{c}(k)$ تقریباً همه جا پیوسته و در صفر مشتق پذیر است. همچنین $\bar{c}(k_0) = \bar{c}_0$. حال مسئله ساده مطرح شده، برای حالتی که چندین محدودیت ورودی داریم تعمیم داده می‌شود. فرض کنید قیمت های خرید $n, \dots, 1, 0, \bar{c}_i, i = 0, 1, \dots, n$ محدودیت داده شده باشند و نیز $\bar{s}_0 = \bar{c}_0$. همچنین با توجه به اینکه تابع $f(x)$ باید در محدودیت تابع چگالی بودن نیز صدق کند. در مجموع $(n + 2)$ محدودیت وجود دارند که تحت آنها باید تابع چگالی بدون مخاطره برای قیمت نهایی دارایی تحت تملک تعیین شود. بنابراین $(n + 1)$ محدودیت دیگر هستند که $n + 1$ معادله دیگر را نتیجه می‌دهد و به‌طور کلی با حل $(n + 2)$ معادله به‌دست آمده ضرایب لاگرانژ $\tilde{\lambda}$ و $\tilde{\beta}_0$ و $\tilde{\beta}_1$ و ... و $\tilde{\beta}_n$ را یافته و بدین ترتیب تابع $f(x)$ به‌دست می‌آید. در عمل برای کاربرد آنچه تاکنون گفته شده‌است تابع چگالی $f(x)$ به‌صورت زیر بیان می‌شود:

$$f(x) = (\tilde{\lambda} + \tilde{\beta}_0 x + \sum_{i=1}^n \tilde{\beta}_i v_i(x))^{-\mu} \quad (20)$$

که $v_i(x)$ نشان دهنده تابع پرداخت i امین وسیله خریداری شده است. همچنین $\mu = \frac{1}{1-\alpha}$ که با توجه به اینکه قبلاً ثابت شد که $1 < \alpha < \frac{3}{2}$ بنابراین باید $\mu > 2$ باشد. که قطعاً می‌توان در حالت‌های خاص با محاسبات عددی جواب مناسب را یافت. با جایگذاری معادله (20) در $(n + 2)$ رابطه داده شده با

۱۱۶ آنتروپی ماکسیم برای اندازه‌های اطلاع تعمیم‌یافته

محدودیت تابع چگالی بودن و معادله (۱۷) و معادلات مختلف به صورت (۱۸) معادلات محدودیت‌های مناسب به دست می‌آیند.

بحث و نتیجه‌گیری

در این مقاله روش آنتروپی ماکسیم برای برخی از اندازه‌های آنتروپی تعمیم‌یافته از قبیل آنتروپی‌های رنی و تی‌سالیس مطرح و ملاحظه شد که توزیع‌های دارای آنتروپی رنی ماکسیم به شکل توزیع‌های توانی هستند. برخی از ویژگی‌های توزیع‌های توانی معرفی و نمایش جدیدی از آنتروپی رنی حاصل گردید. بحث مینیمم اندازه اطلاع رنی مطرح و شکل دیگری از آن به دست آمد. کاربردی اقتصادی از آنتروپی رنی ماکسیم ارائه شد. ضمن یادآوری اندازه واگرایی سیزار برخی از اندازه‌های اطلاع که برای مقادیر مختلف ϕ در این تعریف صدق می‌کنند بیان شد. بحث مینیمم اندازه واگرایی سیزار بررسی و توابع چگالی (احتمال) که با توجه به محدودیت‌های مشخص اندازه‌های اطلاع ذکر شده را مینیمم می‌کنند به دست آمدند.

تقدیر و تشکر

نویسندگان از پیشنهادات داوران محترم که باعث اصلاحات سازنده در مقاله شد کمال تشکر را دارند.

مراجع

صانعی طبس، م. محتشمی برزادران، غ. و امینی، م. ۱۳۸۹، ویژگی‌ها و مشخصه‌هایی از آنتروپی رنی ماکسیم، مجله بررسی‌های آمار رسمی ایران، ۲۱، ۸۷-۷۵.

وحیدیان کامیاد، ع. و بزرگ‌نیا، الف. ۱۳۷۲، مقدمه‌ای بر نظریه کنترل و کنترل بهینه، انتشارات دانشگاه فردوسی مشهد.

دستمرد، ز. محتشمی برزادران، غ. و مقدس زاده بزاز، ب. ۱۳۹۰، تعمیم‌های توزیع‌های گسسته و ویژگی‌های آن‌ها برای اندازه اطلاع، مجله علوم آماری، ۵، ۱۷۸-۱۶۱.

منصوری، ش. پاشا، ع. ۱۳۸۸، تاثیر ضریب همبستگی بر میزان تغییر آنتروپی توزیع توام ماکسیمم آنتروپی، مجله علوم آماری، ۳، ۱۹۶-۱۸۵.

- Bashkirov, A. G. (2004), Maximum Renyi Entropy Principle for Systems with Power-law Hamiltonians, *Physical Review Letters*, **93**, 130601.
- Bashkirov, A. G. (2006), Renyi Entropy as a Statistical Entropy for Complex Systems, *Theoretical and Mathematical Physics A*, **149**, 1573-1559.
- Bercher, J. F. (2008), On Some Entropy Functionals from Renyi Information Divergence, *Information Sciences*, **178**, 2489-2506.
- Brody, D. C., Buckley, I. R. C. and Constantinou, I. (2007), Option Price Calibration from Renyi Entropy, *Physics Letters A*, **366**, 298-.307
- Costa, J. A., Hero, A. O. and Vignat, C. (2002), A Characterization of the Multivariate Distributions Maximizing Renyi Entropy, *IEEE International Symposium on Information Theory, ISIT, Lausanne*, 263.
- Costa, J. A., Hero, A. O. and Vignat, C. (2006), A Geometric Characterization of Maximum Renyi Entropy Distributions, *IEEE International Symposium on Information Theory, ISIT, Lausanne*, 1822-1826.
- Csiszar, I. (1967), On Topological Properties of F-divergences, *Studia Mathematicarum Hungarica*, **2**, 329-339.
- Jaynes, E. T. (1957), Information Theory and Statistical Mechanics, *Physical Reviews*, **106**, 620-630.
- Johnson, O. and Vignat, C. (2007), Some Results Concerning Maximom Renyi Entropy Distributions, *Annales de Institut Henri Poincare (B) Probability and Statistics*, **43**, 339-351.
- Kagan, A. M., Linnik, Y. V. and Rao, C. R. (1973), *Characterization Problems in Mathematical Statistics*, John Wiley, New York.
- Kapur, J. N. (1989), *Maximum Entropy Models in Sciences and Engineering*, John Wiley, New York.
- Kullback, S. and Leibler, R. A. (1951), On Information and Sufficiency, *The Annals of Mathematical Statistics*, **22**, 79-86.
- Kullback, S. (1959), *Information Theory and Statistics*, John Wiley, New York.
- Nagy, A. and Romera, E. (2009), Maximum Renyi Entropy Principle and the Generalized Thomas-Fermi Model, *Physics Letters A*, **373**, 844-846.
- Renyi, A. (1961), On Measures of Entropy and Information, *Proc Berekely Symposium, Statistical Probability*, **1**, 547-561.
- Shannon, C. E. (1948), A Mathematical Theory of Communication, *Bell System Technical*, **27**, 379-423.

آنتروپی ماکسیمم برای اندازه‌های اطلاع تعمیم‌یافته ۱۱۸

Tsallis, C. (1988), Possible Generalizations of Boltzmann-Gibbs Statistics,
Journal of Statistical Physics, **52**, 479-487.