دانشگاه شهید بهشتی

دانشکده فیزیک

پایان‌نامه جهت اخذ درجه کارشناسی ارشد فیزیک

گرایش سیستم‌های پیچیده و دینامیک غیرخطی

مدل‌سازی و تخمین حافظه برای فرایندهای غیرمارکوف وابسته

اساتید راهنما:

دکتر سید علی حسینی

دکتر غلام‌رضا جعفری

استاد مشاور:

دکتر طیب جمالی

نگارنده: حسین مطهری

شهریور ۹۹

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

تشکر و قدردانی

از اساتید محترم راهنما دکتر سید علی حسینی و دکتر غلام‌رضا جعفری و همه دوستان عزیزی که مرا در به ثمر رساندن این پایان‌نامه یاری کردند، به ویژه از دکتر طیب جمالی که در شکل‌گیری این پایان‌نامه نقش شایانی داشته‌اند، صمیمانه تشکر می‌کنم.

چکیده

در دنیای واقعی سیستم‌هایی وجود دارند که دارای رفتار تصادفی هستند و برای مدل‌سازی آن‌ها از فرایندهای تصادفی استفاده می‌کنیم. از آنجایی که هیچ سیستمی کاملا مستقل نیست و با دیگر سیستم‌ها برهمکنش دارد در تحلیل و بررسی این سیستم‌ها بهتر است وابستگی آن‌ها به یکدیگر را در نظر داشت. در کنار این وابستگی یک سیستم می‌تواند به گذشته خودش نیز وابسته باشد یا به عبارت دیگر حافظه‌دار باشد. در فرایندهای تصادفی اگر فرایندی حافظه‌دار باشد به آن فرایند غیرمارکوف گفته می‌شود. به عنوان مثال بازارهای مالی به عنوان نمونه‌ای از سیستم‌های پیچیده هستند که رفتار آن‌ها را می‌توان با فرایندهای تصادفی که به یکدیگر وابسته و غالبا به گذشته خود نیز وابسته هستند مدل‌سازی کرد. با توجه به نقش سیستم‌های پیچیده از قبیل بازارهای مالی در زندگی بشر وجود روش‌هایی برای تخمین حافظه و وابستگی فرایندهای تصادفی به یکدیگر ضروری است. یکی از روش‌های مرسوم برای تخمین حافظه یک فرایند غیرمارکوف(طول مارکوف)، آزمون چپمن-کولموگروف است. هدف از این پایان‌نامه تعمیم آزمون چپمن-کولموگروف به دو فرایند غیرمارکوف وابسته است. در ادامه به منظور نشان دادن کارایی در تخمین طول مارکوف برای دو فرایند غیرمارکوف وابسته نیازمند مدلی هستیم که سری‌های زمانی وابسته با طول‌های مارکوف دلخواه را تولید کند. بدین منظور مدل خودبرگشت را تعمیم می‌دهیم. نتایج نشان می‌دهد که آزمون چپمن-کولموگروف تعمیم یافته با دقت بالایی طول‌های مارکوف را در دو فرایند غیرمارکوف وابسته تخمین می‌زند. در انتها با این روش به بررسی دو رمزارز بیتکوین و اتریوم پرداخته می‌شود و نتایج نشان می‌دهد که نوسانات بزرگ از لحاظ حافظه به یکدیگر وابسته هستند ولی نوسانات کوچک از یکدیگر مستقل هستند.

کلمات کلیدی:

فرایندهای غیر مارکوف وابسته، آزمون چپمن-کولموگروف، حافطه، طول مارکوف، مدل خودبرگشت، رمزارز، بیتکوین، اتریوم

فهرست مطالب

فهرست تصاویر

فهرست جداول

پیشگفتار

در طول تاریخ انسان‌ها همیشه در تلاش بوده‌اند تا محیط پیرامون خود را بهتر بشناسند و رفته رفته این تلاش‌ها باعث پیدایش علوم مختلف شده است. در این میان به دلیل محدود بودن توانایی انسان‌ها و همچنین برخی محدودیت‌های بنیادی باعث شده تا انسان‌ها دست به گسترش مباحثی بزنند که می‌تواند به شکلی بر این محدودیت‌ها غلبه کنند. به عنوان مثال مکانیک آماری یکی از مواردی است که برای بررسی سیستم‌های بس ذره‌ای استفاده می‌شود؛ سیستم‌هایی که حل دقیق آن‌ها برای ما ممکن نیست. یکی دیگر از مباحثی که برای غلبه بر این ناتوانی استفاده می‌شود مبحث فرایندهای تصادفی است.برخی سیستم‌ها به دلیل رفتار شبه تصادفی که دارند نمی‌توان آن‌ها را به شکل دقیق بررسی کرد. اما با استفاده از فرایندهای تصادفی می‌توان رفتار این سیستم‌ها را مدل کرد. فیزیک [۱]، شیمی [۲]، بیولوژی [۳] و پردازش سیگنال [۴] از جمله مواردی هستند که فرایندهای تصادفی به شکل گسترده در آن‌ها استفاده می‌شود. بازارهای مالی یکی دیگر از مواردی است که برای تحلیل و بررسی آن از فرایندهای تصادفی استفاده می‌شود. [۵، ۶]

در دنیای واقعی بررسی سیستم‌ها به شکل مستقل و بدون در نظر گرفتن برهم‌کنش با دیگر سیستم‌ها اطلاعات جامع و کاملی از تحول این سیستم‌ها به ما نمی‌دهد ممکن است سیستمی با سیستم‌های دیگر برهم‌کنشی ضعیف یا حتی قابل صرف نظر کردن داشته باشد اما در حقیقت هیچ سیستمی در واقعیت کاملاً مستقل نیست. در بسیاری از موارد بررسی سیستم‌ها به شکل مستقل می‌تواند از درک درست ساز و کار سیستم مورد مطالعه جلوگیری کند و در صورتی که هدف پیشبینی آینده یک سیستم دینامیکی باشد بررسی آن به شکل مستقل می‌تواند باعث بروز خطا در پیشبینی بشود. از بحران مسکن آمریکا در سال ۲۰۰۸ بارها به عنوان یک مثال مهم از اهمیت اثرگذاری بازارها بر یکدیگر نامبرده شده است، این بحران که از سال ۲۰۰۵ شروع و تا سال ۲۰۰۸ ادامه پیدا کرد، اثری مخرب بر دیگر بازارهای مالی آمریکا و دیگر نقاط دنیا داشت و باعث به وجود آمدن یک رکود اقتصادی در دنیا شد، که نشان دهنده اهمیت ارتباط و اثر گذاری بازارهای مالی بر یکدیگر است. [۷، ۸] بنابراین باید تلاش کرد تا سیستم‌ها را به شکل وابسته با یکدیگر در نظر بگیریم تا بتوان تحول این سیستم‌ها را به شکلی دقیق بررسی کنیم.

نکته مهم دیگر در مورد سیستم‌ها و فرایندهای واقعی وابستگی آن‌ها به حالت‌های گذشته خودشان است که اصطلاحا به آن‌ها سیستم‌ها یا فرایندهای حافظه‌دار یا غیرمارکوف گفته می‌شود. در اینگونه سیستم‌ها حالت‌های گذشته یک سیستم در آینده آن تاثیر گذاز است. تاکنون مدل‌های مختلفی برای سیستم‌های حافظه‌دار ارائه شده است. مثلا مدل خودبرگشت[۱] یکی از ساده‌ترین مدل‌هایی است که می‌توان با آن سری‌های زمانی حافظه‌دار را تولید و بررسی کرد. [۹، ۱۰] مدل‌های بسیار زیادی از این دست وجود دارند که با تعیین پارامترهای آن‌ها می‌توان سری‌های زمانی حافظه‌دار تولید کرد یا با تخمین پارامترها از روی سری‌های زمانی واقعی مانند بازارهای مالی می‌توان سیستم‌ها و فرایندهای واقعی را بررسی کرد. [۹] از جمله مدل‌های دیگری که از آن‌ها در بررسی سری‌های زمانی حافظه‌دار استفاده می‌شود می‌توان به مدل میانگین متحرک خودبرگشت یکپارچه اشاره کرد که به طور گسترده‌ای نیز در تحلیل و پیشبینی بازارهای مالی استفاده شده است[۲]. [۱۱-۱۳].

روش‌های مختلفی برای تخمین حافظه مانند محاسبه خودهمبستگی[۳] وجود دارد. [۱۴] در این میان روش‌های دیگری نیز وجود دارد که می‌توانند تخمینی از حافظه یا طول مارکوف فرایندها به دست آورند، از جمله روش‌هایی که با آن طول مارکوف یک فرایند را تخمین می‌زنند می‌توان به روش χ^2 [۱۵]، روش ویلکاکسون[۴] [۱۶] و آزمون چپمن–کولموگروف[۵] [۱۷-۲۰] اشاره کرد. در این میان آزمون چپمن–کولموگروف به دلیل اینکه با استفاده از توزیع احتمال محاسبه می‌شود می‌توان با تعمیم آن وابستگی یک فرایند به فرایند دیگر را نیز آزمود تا بتوان برای دو فرایند وابسته طول مارکوف و وابستگی را تخمین زد. در این پایان‌نامه آزمون چپمن–کولموگروف را برای دو فرایند غیرمارکوف وابسته تعمیم می‌دهیم تا غیر از تخمین طول حافظه بتوانیم طول وابستگی دو فرایند را تخمین بزنیم.

ابتدا در فصل اول به توضیح مقدمات احتمال و مفاهیم مورد نیاز احتمالات برای تعمیم آزمون چپمن–کولموگروف خواهیم پرداخت. در این فصل مباحث مقدماتی احتمال از جمله تعریف توزیع‌های احتمال توأم و شرطی توضیح داده شده‌اند. در فصل دوم تلاش کردیم تا مفاهیم کلی و مقدمات مورد نیاز از فرایندهای تصادفی را بیان کنیم. در انتهای این فصل فرایندهای غیرمارکوف را تعریف می‌کنیم و روش‌های محاسبه طول حافظه که به آن طول مارکوف می‌گوییم را معرفی می‌کنیم. در فصل سوم به کار خودمان می‌پردازیم و آزمون چپمن–کولموگروف را برای دو فرایند غیرمارکوف وابسته تعمیم خواهیم داد. برای اینکه ببینیم آزمون چپمن–کولموگروف تعمیم داده شده می‌تواند به

[1] Autoregressive model

[2] Autoregressive integrated moving average

[3] Autoorrelation

[4] Wilcoxon test

[5] Chapman-Kolmogorov test

درستی نتیجه بدهد، مدل خودبرگشت را برای تولید دو فرایند غیرمارکوف وابسته تعمیم می‌دهیم و آزمون تعمیم یافته را روی داده‌های این مدل آزمایش می‌کنیم. در فصل ۴ نیز با استفاده از آزمون تعمیم یافته طول مارکوف دو رمز ارز بیتکوین و اتریوم را برای چند بازه زمانی محاسبه می‌کنیم. در نهایت در فصل آخر نیز به جمع‌بندی و نتیجه‌گیری نتایج به دست آمده خواهیم پرداخت.

فصل ۱

مقدمه‌ای بر احتمالات

در این فصل به بیان برخی مفاهیم مورد نیاز از احتمالات خواهیم پرداخت که در ادامه به آن نیاز خواهیم داشت.

۱.۱ احتمال

در نظریه احتمال به هر روندی که بتوان آن را بی‌نهایت بار تکرار کرد و همچنین بتوان یک مجموعه خوش تعریف از تمام نتایج ممکن برای آن تعریف کرد **آزمایش تصادفی** گفته می‌شود. به مجموعه تمام نتایج ممکن یک آزمایش تصادفی **فضای نمونه** گفته می‌شود. به عنوان مثال مسأله پرتاب سکه یک آزمایش تصادفی است و اگر روی سکه را با h و پشت سکه را با t نمایش دهیم، فضای نمونه آن $\Omega = \{h,t\}$ است یا برای آزمایش تصادفی پرتاب تاس فضای نمونه به شکل $\Omega = \{1,2,3,4,5,6\}$ است. معمولاً فضای نمونه را با Ω نشان می‌دهند.

در یک آزمایش تصادفی به هر زیرمجموعه از فضای نمونه **رویداد** گفته می‌شود. اگر نتیجه آزمایش تصادفی عضوی از یک رویداد باشد می‌گوییم آن رویداد رخ داده است. به عنوان مثال اگر دو سکه را پرتاب کنیم فضای نمونه آن به شکل زیر است.

$$\Omega = \{\{h,h\},\{h,t\},\{t,h\},\{t,t\}\}$$

فرض کنید رویداد A رو آمدن سکه اول باشد بنابراین مجموعه A به شکل $A = \{\{h,h\},\{h,t\}\}$ نوشته می‌شود. فضای نمونه Ω را هم می‌توانیم یک رویداد تعبیر کنیم، رویدادی که همیشه رخ می‌دهد و به آن رویداد حتمی

می‌گوییم.

از این پس مجموعه E را مجموعه‌ی همه‌ی رویدادهای یک آزمایش تصادفی در نظر می‌گیریم، بنابراین همه‌ی اعضای E زیر مجموعه‌ی Ω هستند. در ریاضیات به جفت $\{\Omega, E\}$ یک فضای قابل اندازه‌گیری گفته می‌شود که E در واقع جبر سیگما[۱] یا میدان سیگمای فضای نمونه Ω است.

در این ادامه به معرفی تعاریف مختلف احتمال می‌پردازیم.

تعریف کلاسیک: طبق این تعریف احتمال $P(A)$ رویداد A به صورت $P(A) = \frac{N_A}{N}$ تعریف می‌شود که N تعداد همه‌ی نتایج ممکن آزمایش تصادفی و N_A تعداد نتایج مطلوب برای رویداد A است. این تعریف فقط زمانی درست است که هر یک از نتایج ممکن احتمال یکسانی داشته باشند. مثلا در مسأله پرتاب تاس احتمال زوج آمدن تاس برابر با $\frac{3}{6}$ است که با این فرض به دست آمده که آمدن هر وجه تاس احتمال یکسانی دارد. تعریف کلاسیک احتمال تاکنون با انتقادهای زیادی روبرو شده است به عنوان مثال فرض برابر بودن احتمال نتایج باعث سخت شدن تعیین N و N_A می‌شود. همچنین استفاده از این تعریف به مسائل خاصی محدود است. [۲۱]

تعریف فرکانس نسبی[۲]: در این تعریف احتمال رویداد A از رابطه‌ی زیر به‌دست می‌آید.

$$\lim_{n\to\infty} \frac{n_a}{n} \tag{۱.۱}$$

در رابطه بالا n تعداد کل آزمایش‌ها و n_a تعداد دفعاتی است که رویداد A رخ داده است. ایراد این تعریف این است که در آزمایش‌های فیزیکی عددهای n_a و n ممکن است خیلی بزرگ باشد اما هیچ‌گاه بی‌نهایت نیست. بنابراین در واقعیت نمی‌توان از آزمایش حد ۱.۱ را محاسبه کرد و تنها می‌توانیم فرض کنیم که نسبت $\frac{n_A}{n}$ برابر حد ۱.۱ است.

اصول موضوعه احتمال: این اصول موضوعه که با نام **اصول کولموگروف**[۳] نیز شناخته می‌شوند شامل سه اصل زیر هستند.

۱) احتمال رویداد A عددی مثبت است که به رویداد A نسبت داده می‌شود.

$$P(A) \geq 0 \tag{۲.۱}$$

[1] σ-algebra
[2] Relative frequency
[3] Kolmogorov axioms

۲) به رویدادی که در هر بار انجام آزمایش رخ بدهد رویداد قطعی گفته می‌شود. اگر S یک رویداد قطعی باشد احتمال این رویداد برابر ۱ است.

$$P(S) = 1 \tag{۳.۱}$$

۳) اگر رویداد A و B دو رویداد ناسازگار باشند به این معنی که اشتراک آن‌ها مجموعه‌ی تهی است. داریم:

$$P(A \cup B) = P(A) + P(B) \tag{۴.۱}$$

اصول موضوعه احتمال به عنوان یک تعریف تئوری به دو تعریف دیگر ترجیح داده می‌شوند هرچند تعریف فرکانس نسبی برای کاربردهای احتمال در مسائل واقعی ضروری است. در واقع تعریف فرکانس نسبی احتمال $P(A)$ را به دفعات مشاهده یک رویداد مرتبط می‌سازد. [۲۲]

سه مفهوم فضای نمونه، رویداد و احتمال در کنار هم، فضایی را تشکیل می‌دهند که به آن فضای احتمال گفته می‌شود. فضای احتمال در ریاضی به شکل $\{\Omega, E, P\}$ نوشته می‌شود. در ریاضی به جفت $\{\Omega, E\}$ یک فضای قابل اندازه‌گیری [۴] گفته می‌شود و در ریاضیات می‌توان برای یک فضای قابل اندازه‌گیری یک سنجه تعریف کرد که طبق تعریف، سنجه به هر عضو از E که همان رویدادها هستند یک عدد نسبت می‌دهد. سنجه‌ای که از اصول موضوعه احتمال پیروی کند را احتمال می‌گوییم و آن را P نشان می‌دهیم.

۲.۱ احتمال توأم و شرطی

در احتمالات اگر بخواهیم دو یا چند رویداد را به صورت مشترک بررسی کنیم باید از مفهومی به نام احتمال توأم [۵] استفاده کنیم. ممکن است این رویدادها مربوط به آزمایش‌های تصادفی متفاوتی باشند اما به شکلی به هم وابسته‌اند. مثلا در بسیاری از موارد علاقه‌مند هستیم که احتمال رخ دادن دو یا چند رویداد را که به هم وابسته هستند به صورت همزمان بررسی کنیم که برای این کار باید از احتمال توأم استفاده کنیم.

تعریف احتمال توأم به شکل زیر است که A_iها رویدادهایی هستند که ممکن است متعلق به آزمایش‌های

[4] Measurable space

[5] Joint probability

تصادفی متفاوتی باشند.

$$P(A_1, A_2, A_3, \dots A_n) = P(A_1 \cap A_2 \cap \dots A_n) \tag{۵.۱}$$

به عنوان مثالی برای استفاده از احتمال توأم فرض کنید رویداد A زوج آمدن یک تاس و رویداد B عددی بزرگ‌تر از ۳ آمدن آن تاس باشد. بنابراین رویداد $A = \{2, 4, 6\}$ و رویداد $B = \{4, 5, 6\}$ و اشتراک این دو برابر با $A \cap B = \{4, 6\}$ است. بنابراین احتمال توأم A و B، احتمال ۴ یا ۶ آمدن تاس را بیان می‌کند. اگر دو رویداد به طور مشترک رخ ندهند یعنی اشتراکی با هم ندارند و بنابراین احتمال توأم آن‌ها صفر است.

احتمال شرطی[6] مفهوم دیگری است که بیان کننده احتمال رخ دادن یک یا چند رویداد به شرطی است که یک یا چند رویداد دیگر رخ داده باشند یا بتوانیم فرض کنیم که رخ داده‌اند. احتمال شرطی رویدادهای $A_1, A_2, \dots, A_n$ به شرط رویدادهای $B_1, B_2, \dots, B_m$ را با $P(A_1, A_2, \dots A_n \mid B_1, B_2, \dots B_m)$ نشان می‌دهند.

احتمال شرطی نیز مانند احتمال توأم برای بررسی رویدادهایی است که به نوعی به هم وابسته‌اند مثلاً فرض کنید در یک روز آفتابی احتمال رخ دادن یک تصادف رانندگی در یک خیابان برابر ۰٫۰۵ و در یک روز بارانی برابر با ۰٫۱۵ باشد. در هر دو حالت، احتمال رخ دادن تصادف مورد نظر ما است، اما این احتمال وابسته به رویداد دیگری است که در اینجا شرایط آب و هوا است.

طبق تعریف کولموگروف رابطه‌ی احتمال شرطی و احتمال توأم به شکل زیر تعریف می‌شود.

$$P(A \mid B) = \frac{P(A, B)}{P(B)} \tag{۶.۱}$$

اگر رابطه بالا را در P(B) ضرب کنیم داریم:

$$P(A, B) = P(B)P(A \mid B) \tag{۷.۱}$$

به عبارت دیگر رابطه بالا بیان می‌کند که احتمال اینکه A و B هر دو رخ بدهند برابر است با احتمال رخ دادن B ضرب در احتمال رخ دادن A به شرطی که B رخ داده باشد. شکل کلی رابطه ۷.۱ به صورت زیر است که به آن

[6] Conditional probability

قانون ضرب هم گفته می‌شود.

$$P(A_1 \cap A_2 \cap \dots A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1, A_2) \dots P(A_n \mid A_1 \dots A_{n-1}) \quad (۸.۱)$$

حالت خاصی وجود دارد که دو رویداد، مثلاً A و B به هم وابسته نیستند و اصطلاحاً به آن‌ها دو رویداد مستقل گفته می‌شود، در این صورت احتمال توأم آن‌ها به صورت ضرب احتمال تک تک رویدادها نوشته می‌شود.

$$P(A, B) = P(A \cap B) = P(A)P(B) \quad (۹.۱)$$

رابطه بالا درواقع تعریف دو رویداد مستقل نیز هست یعنی اگر رابطه بالا را بتوان برای دو رویداد نوشت آن دو رویداد حتما مستقل از هم هستند. اگر رابطه ۹.۱ را در رابطه ۶.۱ جایگذاری کنیم، داریم

$$P(A \mid B) = \frac{P(A)P(B)}{P(B)} = P(A)$$

که بیانگر این است که رویداد A وابستگی به B ندارد. واضح است که این رابطه را برای وابستگی رویداد B به A نیز می‌توان نوشت. [۲۳]

۳.۱ متغیر تصادفی

همان‌طور که در بخش ۱.۱ گفته شد هر آزمایش تصادفی دارای یک فضای نمونه است که مجموعه‌ای از تمام نتایج ممکن آزمایش تصادفی است. متغیر تصادفی عددی است که به هر یک از نتایج یک آزمایش تصادفی نسبت می‌دهیم به عنوان مثال در یک آزمایش پرتاب سکه می‌توانیم متغیر تصادفی را به این شکل تعریف کنیم که اگر روی سکه آمد مقدار متغیر برابر ۱ و اگر پشت سکه آمد مقدار متغیر برابر صفر باشد. به عبارت دیگر متغیر تصادفی تابعی است که دامنه‌ی آن فضای نمونه آزمایش تصادفی و برد آن یک فضای قابل اندازه‌گیری است. اگر x متغیر تصادفی

و r نتیجه‌ی آزمایش تصادفی باشد می‌توان گفت:

$$\mathbf{x} : \Omega \to S$$
$$\mathbf{x}(r \in \Omega) = X \in S$$

در تعریف بالا S یک فضای قابل اندازی‌گیری است که در بسیاری موارد $S = \mathbb{R}$ است. اگر $X \in S$ باشد، احتمال اینکه $\mathbf{x} = X$ بشود به شکل زیر تعریف می‌شود.

$$P(X \in S) = P(\{r \in \Omega \mid \mathbf{x}(r) = X\})$$

در برخی از شرایط نیاز است که بدانیم احتمال اینکه متغیر تصادفی $\mathbf{x}$ مقداری کمتر از عدد X بگیرد چقدر است، یا احتمال اینکه $\mathbf{x}$ بین X_1 و X_2 باشد چقدر است. برای یافتن این احتمال باید نتایجی را یافت که متغیر تصادفی $\mathbf{x}$ برای آن‌ها بین X_1 و X_2 باشد، احتمال این رویداد همان احتمالی است که $\mathbf{x}$ بین X_1 و X_2 باشد. به عبارت دیگر:

$$P(X_1 \leq \mathbf{x} \leq X_2) = P(\{r \in \Omega \mid X_1 \leq \mathbf{x}(r) \leq X_2\})$$

۱.۳.۱ متغیر تصادفی پیوسته

تا اینجا تنها متغیرهای تصادفی با فضای نمونه گسسته را بررسی کرده‌ایم و همان‌طور که گفته شد P تابعی است که احتمال رویدادها را مشخص می‌کند معمولاً به P **تابع توزیع احتمال**[۷] گفته می‌شود، اما متغیرهای تصادفی پیوسته نیز وجود دارند که فضای نمونه آن‌ها ناشمارا است به عبارت دیگر مقادیر یک متغیر تصادفی پیوسته می‌تواند هر مقداری از یک بازه‌ی پیوسته باشد. در این شرایط چون فضای نمونه یک بازه پیوسته است و در نتیجه تعداد ناشمارایی رویداد خواهیم داشت باید احتمال رویدادها را به شکل دیگری بررسی کنیم.

فرض کنید $f(x)$ یک تابع مثبت باشد که در بازه $(-\infty, +\infty)$ تعریف شده باشد. همچنین احتمال اینکه مقدار X عضوی از مجموعه A باشد را بتوان با انتگرال f روی مجموعه A به دست آورد، در این صورت x را یک متغیر

[7] Probability distribution function

تصادفی پیوسته می‌گوییم. به عبارت دیگر

$$P\{x \in A\} = \int_A f(x)dx \tag{۱۰.۱}$$

تابع f را **تابع چگالی احتمال**[۸] متغیر تصادفی x گوییم. از آنجایی که x حتماً باید مقدار داشته باشد، بنابراین

$$P\{x \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx = 1 \tag{۱۱.۱}$$

که بیانگر آن است که جمع روی احتمال همه نتایج ممکن برابر با یک است، به رابطه بالا شرط نرمال بودن می‌گویند که شرطی لازم برای هر توزیع احتمال است.

اگر در رابطه ۱۰.۱ مجموعه A می‌تواند بازه $[a, b]$ باشد بنابراین

$$P\{a \leq x \leq b\} = \int_a^b f(x)dx \tag{۱۲.۱}$$

که احتمال اینکه x مقداری بین a تا b داشته باشد را بیان می‌کند. در رابطه ۱۲.۱ اگر $a = b$ داریم

$$P\{x = a\} = \int_a^a f(x)dx = 0$$

که به این معنی است که احتمال اینکه متغیر تصادفی یک مقدار معین داشته باشد صفر است. از آنجایی که انتگرال بالا به معنی احتمال است بنابراین تابع چگالی احتمال یک تابع بی بعد مانند توزیع احتمال نیست و بعدی عکس بعد متغیر تصادفی دارد.

کمیت مهم دیگر **تابع توزیع تجمعی**[۹] است که آن را با $F(y)$ نشان می‌دهیم و رابطه آن با تابع چگالی احتمال به شکل زیر است

$$F(y) = P\{x \in (-\infty, a]\} = \int_{-\infty}^{y} f(x)dx \tag{۱۳.۱}$$

[8] Probability density function

[9] Cumulative distribution function

اگر از دو طرف رابطه بالا مشتق بگیریم داریم

$$\frac{d}{dy}F(y) = f(y) \tag{۱۴.۱}$$

که نشان دهنده این است که تابع چگالی احتمال مشتق توزیع تجمعی است.

از رابطه ۱۲.۱ می‌توانیم به تفسیر دیگری از چگالی احتمال برسیم، اگر بازه انتگرال $(a - \frac{\varepsilon}{2}, a + \frac{\varepsilon}{2})$ باشد و ε مقدار کوچکی داشته باشد، با استفاده از بسط تیلور می‌توانیم بنویسیم:

$$P\left\{a - \frac{\varepsilon}{2} \leq x \leq a + \frac{\varepsilon}{2}\right\} = \int_{a-\varepsilon/2}^{a+\varepsilon/2} f(x)dx \approx \varepsilon f(a) \tag{۱۵.۱}$$

به عبارت دیگر احتمال اینکه x مقداری در بازه‌ای به طول ε اطراف a داشته باشد تقریباً برابر $\varepsilon f(a)$ است، مشخص است که $f(a)$ معیاری است از احتمال اینکه متغیر تصادفی در نزدیکی a باشد.

یکی از روابط مهم در خصوص توزیع و چگالی احتمال، رابطه توزیع مارجینال[10] است که به شکل زیر بیان می‌شود.

$$\begin{aligned} &P(x_1, \dots, x_{i-1}) = \textstyle\sum_i P(x_1, \dots, x_i) && \text{برای متغیر تصادفی گسسته} \\ &f(x_1, \dots, x_{n-1}) = \textstyle\int_{-\infty}^{\infty} f(x_1, \dots, x_n)dx_n && \text{برای متغیر تصادفی پیوسته} \end{aligned} \tag{۱۶.۱}$$

این رابطه در عین سادگی رابطه‌ای مهم در احتمالات است، در فصل بعد از این رابطه استفاده خواهیم کرد.

۴.۱ خواص توزیع احتمال

در این بخش برخی از کمیت‌ها مانند مقادیر چشم‌داشتی و انباشتک‌ها که می‌توان از تابع توزیع احتمال به دست آورد خواهیم پرداخت.

[10] Marginal distribution

۱.۴.۱ مقادیر چشم‌داشتی و ممان

یکی از مهم‌ترین مفاهیم در نظریه احتمالات، مقدار چشم‌داشتی یک متغیر تصادفی است. برای یک متغیر تصادفی گسسته مقدار چشم‌داشتی یا همان میانگین یک متغیر تصادفی از رابطه زیر به دست می‌آید و آن را با $<x>$ نشان می‌دهند.

$$<x>=\begin{cases}\sum_i x_i P(x_i) & \text{برای متغیر تصادفی گسسته}\\ \int x f(x)dx & \text{برای متغیر تصادفی پیوسته}\end{cases} \quad (۱۷.۱)$$

به عبارت دیگر چشم‌داشتی x یک میانگین وزن دار از تمام مقادیر ممکن x است. مثلاً در شرایطی که توزیع احتمال یکنواخت است به این معنی که احتمال همه نتایج با یکدیگر یکسان است مقدار چشم‌داشتی همان رابطه معمولی برای میانگین است.

در بسیاری از مواقع نیاز است که چشم‌داشتی تابعی مثلا $g(x)$ از یک متغیر تصادفی را بدانیم در این صورت به شکل زیر این مقادیر نیز به دست می‌آیند.

$$<g(x)>=\begin{cases}\sum_i g(x_i) P(x_i) & \text{برای متغیر تصادفی گسسته}\\ \int g(x) f(x)dx & \text{برای متغیر تصادفی پیوسته}\end{cases} \quad (۱۸.۱)$$

به چشم‌داشتی توابع به شکل $g(x)=x^k$ که $k=0,1,\dots,n$ است، ممان[11] kاُم گفته می‌شود که از مهمترین مقادیر چشم‌داشتی در نظریه احتمالات هستند. ممان صفرم یعنی زمانی که $k=0$ باشد معادله ۱۸.۱ تبدیل به معادله ۱۱.۱ می‌شود که به معنی جمع احتمال همه نتایج ممکن است و به دلیل نرمال بودن توزیع احتمال برابر با یک است. ممان اول نیز همان رابطه ۱۷.۱ است.

۲.۴.۱ توابع مولد ممان

همان‌طور که در بخش قبل گفته شد ممان‌ها کمیت‌های مهمی هستند، یک توزیع احتمال با استفاده از ممان‌هایش توصیف می‌شود. در این بخش تابعی را معرفی می‌کنیم که با استفاده از آن می‌توان ممان‌های توزیع احتمال یک

[11] Moment

متغیر تصادفی را به دست آورد که به آن تابع مولد ممان[۱۲] گفته می‌شود. با توجه به بخش قبل برای متغیر تصادفی x اگر $g(x) = e^{tx}$ که t یک عدد حقیقی است به تابع زیر تابع مولد ممان گفته می‌شود.

$$M_x(t) = < e^{tx} > = \begin{cases} \sum_i e^{tx_i} P(x_i) & \text{برای متغیر تصادفی گسسته} \\ \int e^{tx} f(x) dx & \text{برای متغیر تصادفی پیوسته} \end{cases} \qquad (۱۹.۱)$$

از آنجایی که به ازای $t = 0$ تابع مولد ممان مقداری برابر با یک دارد پس در $t = 0$ این تابع همیشه با معنی است، به ازای بقیه مقادیر t به شرطی که x متناهی باشد این تابع با معنی است. همان‌طور که از نام تابع مولد ممان مشخص است از این تابع می‌توان برای به دست آوردن ممان‌های یک توزیع احتمال استفاده کرد، برای این کار باید $M_x(x)$ را حول $t = 0$ بسط دهیم. داریم

$$< e^{tx} > = < 1 + tx + \frac{t^2 x^2}{2!} + \cdots > = 1 + E < x > t + E < x^2 > \frac{t^2}{2!} + \cdots \qquad (۲۰.۱)$$

با توجه به بسط تابع مولد ممان می‌توان این‌طور برداشت کرد که ممان‌های یک توزیع احتمال با مشتق تابع مولد ممان در $t = 0$ برابر است. یعنی

$$< x^n > = \left. \frac{d^n M_x(t)}{dt^n} \right|_{t=0} \qquad (۲۱.۱)$$

به این صورت با مشتق گیری از تابع مولد ممان می‌توان ممان‌های یک توزیع را به دست آورد. نکته مهم این است که اگر تابع مولد ممان دو متغیر تصادفی با هم برابر باشد بنابراین توزیع احتمال این دو متغیر تصادفی نیز با یکدیگر برابر است.

[12] Moment generating function

۳.۴.۱ تابع مشخصه

تعریف تابع مشخصه[۱۳] به شکل زیر است.

$$C_x(t) = < e^{itx} > = \begin{cases} \sum_j e^{itx_j} P(x_j) & \text{برای متغیر تصادفی گسسته} \\ \int e^{ix} f(x) dx & \text{برای متغیر تصادفی پیوسته} \end{cases} \qquad (۲۲.۱)$$

تابع مشخصه یک متغیر تصادفی درواقع به نوعی همان تابع مولد ممان است درواقع $M_x(it)$ دقیقاً همان تابع مشخصه است. همان‌طور که تعریف تابع مشخصه و تابع مولد ممان مشابه یکدیگر است موارد استفاده و کاربردهای این دو نیز به یکدیگر نزدیک است. ویژگی‌های که برای تابع مولد ممان گفته شده را می‌توان با استفاده از $C_x(t) = M_x(it)$ برای تابع مشخصه به دست آورد. به عنوان مثال ممان‌های x بر حسب مشتق‌های تابع مشخصه $C_x(t)$ را می‌توان با رابطه زیر به دست آورد.

$$< x^n > = (-i)^n C_x^{(n)}(0) = \mu_n \qquad (۲۳.۱)$$

که در رابطه بالا

$$C_x^{(n)}(0) = \left. \frac{d^n C_x(t)}{d^n t} \right|_{t=0}$$

این که برای توصیف یک متغیر تصادفی از تابع مولد ممان یا تابع مشخصه استفاده کنیم تا حدی یک موضوع شخصی است. اگرچه استفاده از تابع مشخصه مزایایی دارد، مهمترین آن جایگزینی تابع نمایی e^{tx} در تابع مولد ممان با تابع نوسانی مختلط e^{itx} در تابع مشخصه است. مزیت این جایگزینی این است که در تابع مولد ممان باید مراقب واگرا شدن این تابع باشیم اما برای تابع مشخصه به دلیل داشتن رفتار نوسانی نگرانی در این مورد وجود ندارد. علاوه بر این زمانی که x یک متغیر تصادفی پیوسته است رابطه ۲۲.۱ با تبدیل فوریه تابع چگالی احتمال $f(x)$ مرتبط است. بنابراین می‌توان چگالی احتمال $f(x)$ را با استفاده از تبدیل معکوس فوریه از تابع مشخصه به دست آورد.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_x(t) e^{-itx} dt \qquad (۲۴.۱)$$

[13] Characteristic function

۴.۴.۱ تابع مولد انباشتک

انباشتک‌ها[۱۴] از مهمترین کمیت‌های مرتبط با یک توزیع احتمال هستند و می‌توان آن‌ها را بر حسب ممان‌های یک توزیع احتمال به دست آورد. بهترین راه برای به دست آوردن انباشتک‌ها استفاده از تابع مولد انباشتک[۱۵] است که آن را با $K_x(t)$ نشان می‌دهیم. طبق تعریف تابع مولد انباشتک برابر است با لگاریتم طبیعی تابع مولد ممان و اگر $K_x(t)$ را به صورت یک سری توانی از t بسط دهیم ضریب جمله $\frac{t^k}{k!}$ انباشتک kاُم تابع چگالی احتمال $f(x)$ است. یعنی

$$K_x(t) = \ln M_x(t) \equiv \kappa_1 t + \kappa_2 \frac{t^2}{2!} + \kappa_3 \frac{t^3}{3!} + \cdots \qquad (۲۵.۱)$$

می‌توانیم انباشتک‌های یک توزیع را با مشتق گرفتن از تابع مولد انباشتک نسبت به t به دست آوریم

$$\frac{dK_x}{dt} = \frac{1}{M_x}\frac{dM_x}{dt} \Rightarrow M_x \frac{dK_x}{dt} = \frac{dM_x}{dt} \qquad (۲۶.۱)$$

با استفاده از ۲۰.۱ و ۲۵.۱ می‌توانیم بنویسیم

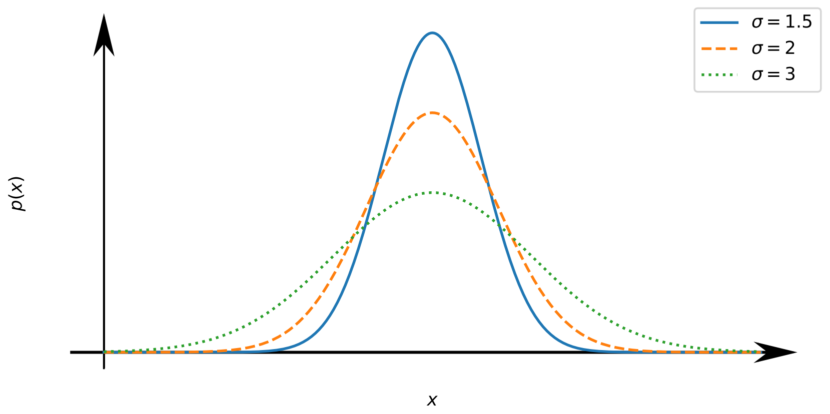
$$(\kappa_1 + \kappa_2 t + \kappa_3 \frac{t^2}{2!} + \cdots)(1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \cdots) = (\mu_1 + \mu_2 t + \mu_3 \frac{t^2}{2!} + \cdots) \qquad (۲۷.۱)$$

با برابر قرار دادن ضرایب توان‌های t در دو طرف تساوی می‌توانیم بنوسیم

$$\begin{aligned}
\mu_1 &= \kappa_1 \\
\mu_2 &= \kappa_2 + \kappa_1 \mu_1 \\
\mu_3 &= \kappa_3 + 2\kappa_2 \mu_1 + \kappa_1 \mu_2 \\
&\vdots \\
\mu_k &= \kappa_k + {}^{k-1}C_1 \kappa_{k-1}\mu_1 + \cdots + {}^{k-1}C_r \kappa_{k-r}\mu_r + \cdots + \kappa_1 \mu_{k-1}
\end{aligned}$$

[14] Cumulant

[15] Cumulant generating function

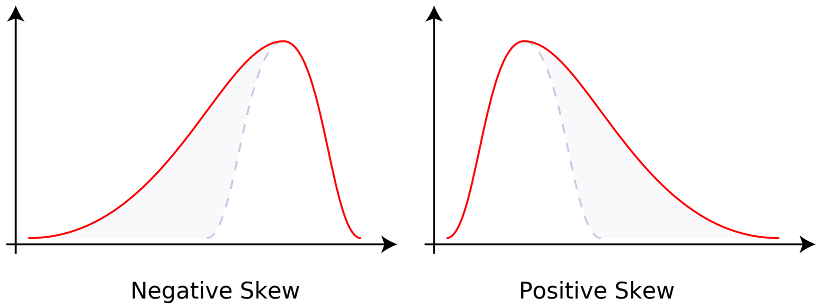


شکل ۱.۱: اثر انحراف معیار بر توزیع نرمال

با حل این معادلات می‌توانیم انباشتک‌ها را بر حسب ممان‌ها به دست آوریم. چهار ممان اول به صورت زیر به دست می‌آیند.

$$
\begin{aligned}
\kappa_1 &= \mu_1 \\
\kappa_2 &= \mu_2 - \mu_1^2 = v_2 \\
\kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = v_3 \\
\kappa_4 &= \mu_4 - 4\mu_3\mu_1 + 12\mu_2\mu_1^2 - 3\mu_2^2 - 6\mu_1^4 = v_4 - 3v_2^2
\end{aligned}
\tag{۲۸.۱}
$$

همان‌طور که پیداست انباشتک اول درواقع همان میانگین توزیع است. به انباشتک دوم واریانس گفته می‌شود همچنین به مجذور واریانس نیز انحراف معیار گفته می‌شود که معمولا آن را با σ نشان می‌دهند. واریانس معیاری از پراکندگی مقادیر متغیر تصادفی x حول میانگین هستند. انباشتک سوم چولگی نام دارد و معیاری از عدم تقارن چگالی احتمال حول میانگین است واگر چولگی برابر با صفر باشد توزیع حول میانگین تقارن کامل دارد. در شکل زیر اثر چولگی مثبت و منفی نشان داده شده است.



شکل ۲.۱: اثر چولگی بر توزیع احتمال [۲۴]

انباشتک چهارم کشیدگی یا برجستگی نام دارد که توصیف کننده میزان مسطح بودن یا قله‌ای بودن یک توزیع احتمال است. هرچه کشیدگی یک توزیع مقدار بزرگ‌تری داشته باشد، آن توزیع قله‌ای‌تر و دنباله پهن‌تری دارد.

۵.۱ توزیع‌های احتمال پرکاربرد

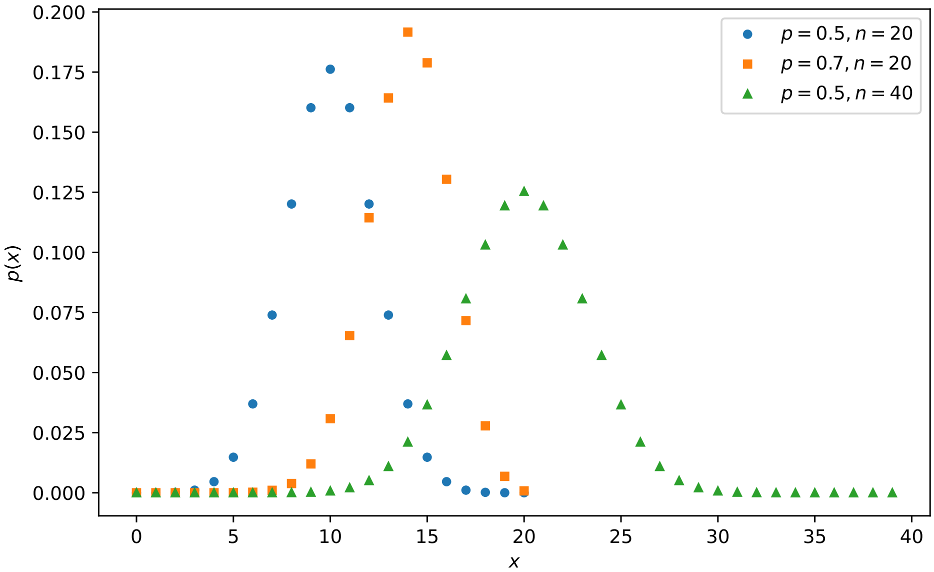
در این بخش به معرفی چند توزیع احتمال پرکاربرد در احتمالات می‌پردازیم.

۱.۵.۱ توزیع دو جمله‌ای

یکی از مهمترین توزیع‌های احتمال گسسته توزیع احتمال دو جمله‌ای [۱۶] است. این توزیع شامل تعدادی مثلاً n بار آزمایش مستقل با دو نتیجه ممکن مثلاً A و B است که این دو نتیجه اشتراکی با هم ندارند. به عنوان مثال این دو نتیجه می‌توانند T یا H آمدن سکه باشد یا موفقیت یا عدم موفقیت یک آزمایش تصادفی باشند. فرض کنید احتمال موفقیت یک آزمایش تصادفی برابر p باشد و احتمال عدم موفقیت برابر $q = 1 - p$ باشد در این صورت اگر n بار آزمایش را مستقل از هم تکرار کنیم می‌توانیم متغیر تصادفی x را برابر تعداد بارهایی که آزمایش موفق است تعریف کنیم. حال احتمال اینکه تعداد x بار آزمایش موفق باشد و $n - x$ بار ناموفق باشد برابر است با

$$\mathrm{P}(x; n, p) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1 - p)^{n-x} \tag{۲۹.۱}$$

[16] Binomial distribution



شکل ۳.۱: چند نمونه از توزیع دو جمله‌ای

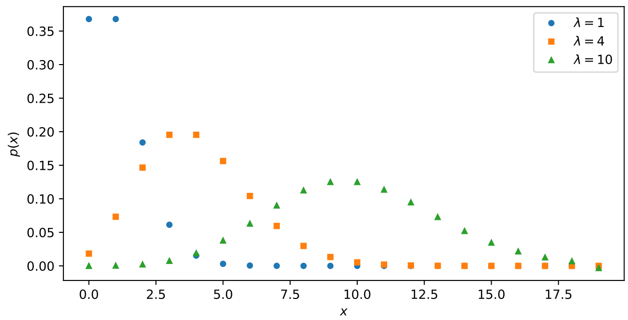
برخی از ویژگی‌های توزیع دو جمله‌ای در جدول زیر آمده است.

جدول ۱.۱: برخی از ویژگی‌های توزیع دو جمله‌ای

میانگین	np
واریانس	npq
چولگی	$\frac{q-p}{\sqrt{npq}}$
کشیدگی	$\frac{1-6pq}{npq}$

۲.۵.۱ توزیع پواسون

توزیع پواسون [17] برای شرایطی است که تعداد دفعات انجام آزمایش زیاد باشد یعنی زمانی که $n \to \infty$، مثلاً فرض کنید که مسأله مورد نظر تعداد تماس‌هایی است که در یک بازه زمانی با یک مرکز تلفن می‌شود. با در نظر گرفتن یک نرخ رخداد λ در واحد زمان، توزیع پواسون $P(x)$ احتمال اینکه دقیقا x بار رویداد در واحد زمان رخ بدهد را نشان می‌دهد. [۲۵]



شکل ۴.۱: چند نمونه از توزیع پواسون

توزیع پواسون درواقع حالت حدی توزیع دو جمله‌ای است زمانی که $n \to \infty$ و احتمال اینکه رویداد رخ بدهد نزدیک به صفر باشد یعنی $p \to 0$، به شرطی که حاصل ضرب np مقداری متناهی λ باشد. شکل ریاضی توزیع پواسون به شکل زیر است.

$$f(k;\lambda) = \mathrm{P}(x = k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{۳۰.۱}$$

برخی از ویژگی‌های توزیع پواسون در ادامه آمده است.

جدول ۲.۱: برخی از ویژگی‌های توزیع پواسون

میانگین	λ	واریانس	λ
چولگی	$\lambda^{-1/2}$	کشیدگی	λ^{-1}

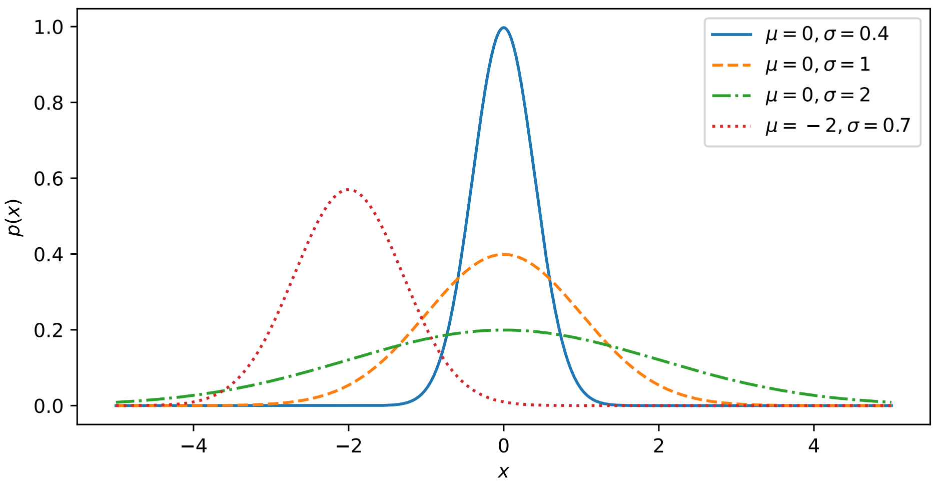
[17] Poisson distribution

۳.۵.۱ توزیع گاوسی

توزیع گاوسی[۱۸] که به آن توزیع نرمال[۱۹] هم گفته می‌شود مهمترین توزیع پیوسته است. دلیل اهمیت توزیع نرمال این است که تعداد خیلی زیادی از متغیرهای تصادفی مورد مطالعه در همه زمینه‌های علمی به طور دقیق یا تقریبی با توزیع گاوسی توصیف می‌شوند. توزیع گاوسی به شکل زیر بیان می‌شود.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{۳۱.۱}$$

در رابطه بالا μ میانگین و σ انحراف معیار هستند و ضریب $\frac{1}{\sigma\sqrt{2\pi}}$ را ضریب نرمال سازی می‌گویند که از شرط نرمال بودن ۱۱.۱ به دست می‌آید. توزیع گاوسی را نیز می‌توان یکی از حالت‌های حدی توزیع دو جمله‌ای در نظر گرفت. توزیع گاوسی در شرایطی به دست می‌آید که $n \to \infty$ و همچنین احتمال p نیز مقداری متناهی داشته باشد و در نتیجه $np \to \infty$.



شکل ۵.۱: چند نمونه از توزیع گاوسی

حدود 68.2% از اعدادی که از توزیع نرمال می‌آیند در بازه $[\mu-\sigma, \mu+\sigma]$ قرار می‌گیرند. یعنی در فاصله‌ای کمتر یا برابر با انحراف معیار نسبت به میانگین قرار می‌گیرند و تقریبا 95.4% از اعداد در بازه $[\mu-2\sigma, \mu+2\sigma]$

[18] Gaussion distribution

[19] Normal distribution

قرار می‌گیرند. یکی دیگر از ویژگی‌های مهم توزیع گاوسی صفر بودن ممان‌های سوم و بالاتر است بنابراین توزیع نرمال یک توزیع متقارن حول میانگین است.

فصل ۲

فرایندهای تصادفی

در فصل قبل به معرفی و بررسی متغیرهای تصادفی پرداختیم. در این فصل قصد داریم یک گام به جلو برداریم و به بررسی فرایندهای تصادفی[1] بپردازیم. از فصل قبل می‌دانیم که x یک متغیر تصادفی با توزیع احتمال $P(x)$ است. فرایند تصادفی به دنباله‌ای از متغیرهای تصادفی گفته می‌شود مانند $x_{t_1}, x_{t_2}, \ldots, x_{t_n}, \ldots$ که پایین‌نویس‌های $\{t_i\}$ می‌توانند نشان دهنده زمان باشند. اما در حالت کلی مجموعه‌ای گسسته یا پیوسته است و لزوما معنای زمان نمی‌دهد. در ادامه مطلب از $\{t_i\}$ به عنوان زمان استفاده خواهد شد اما محاسبات انجام شده در حالت کلی نیز درست است. در هر نقطه زمانی t_i (در حالت پیوسته از t بدون پایین‌نویس استفاده خواهیم کرد) توزیع احتمال می‌تواند با زمان‌های دیگر متفاوت باشد. این وابستگی زمانی را می‌توان به شکل $P(x,t)$ برای توزیع احتمال و به شکل $p(x,t)$ برای چگالی احتمال نشان داد. اگر توزیع احتمال توأم یا چگالی احتمال توأم یک فرایند تصادفی یعنی $p(x_1,t_1;x_2,t_2;\ldots x_n,t_n,\ldots)$ را بدانیم، درواقع همه چیز را در مورد این فرایند تصادفی می‌دانیم چون می‌توانیم احتمال تمام مسیرهای ممکن برای فرایند تصادفی را محاسبه کنیم. [۲۶] در حالت کلی برای یک فرایند تصادفی می‌توانیم بنویسیم.

$$\begin{aligned} &p(x_n,t_n;x_{n-1},t_{n-1}\cdots;x_0,t_0)\\ &\quad = p(x_n,t_n \mid x_{n-1},t_{n-1}\cdots;x_0,t_0)\ldots p(x_2,t_2 \mid x_1,t_1;x_0,t_0)p(x_1,t_1 \mid x_0,t_0)p(x_0,t_0) \end{aligned} \tag{۱.۲}$$

[1] Stochastic process

همان‌طور که از رابطه بالا مشخص است در اکثر موارد نمی‌توان این توزیع احتمال توأم را به دست آورد و در حالت کلی بررسی فرایندهای تصادفی به این شکل کاری بسیار دشوار است. اما تا کنون حالت‌های خاص و مدل‌های مختلفی برای فرایندهای تصادفی ارائه شده است مانند فرایندهای مارکوف[۲]، یا مدل‌های پنهان مارکوف[۳] که هر کدام استفاده‌های مختلف و گسترده‌ای دارند. در شرایط تجربی علاقه‌مندیم که در صورت امکان مسائل خود را با مدل‌های موجود تخمین بزنیم بنابراین به ابزاری نیاز داریم تا بتوانیم امکان تخمین زدن مسائل با مدل‌های تصادفی را بررسی کنیم.

یکی ویژگی‌های مهم فرایندهای تصادفی مانا بودن یا نبودن آن‌ها است. اگر توزیع n نقطه‌ای یک فرایند تصادفی به ازای nهای مختلف وابستگی به زمان نداشته باشد، یعنی:

$$p\left(x_{t_1+\tau}, \ldots, x_{t_n+\tau}\right) = p\left(x_{t_1}, \ldots, x_{t_n}\right) \quad (۲.۲)$$

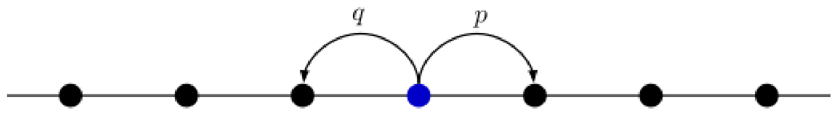
توزیع احتمال در سمت راست و چپ رابطه بالا به اندازه τ اختلاف زمانی دارند. اگر رابطه بالا برای یک فرایند تصادفی برقرار باشد آن فرایند مانا[۴] است. حالت‌های ضعیف‌تری برای مانا بودن یک فرایند وجود دارد مثلاً اگر میانگین و واریانس یک فرایند در زمان تغییر نکند به آن، فرایند مانای ضعیف گفته می‌شود. [۲۶]

۱.۲ ولگشت تصادفی

یکی از ساده‌ترین مثال‌ها برای فرایندهای تصادفی مسأله ولگشت تصادفی[۵] است. ذره‌ای را فرض کنید در هر زمان به شکل کاملاً تصادفی بر روی محور x قدمی به سمت چپ یا راست بر می‌دارد. در حالت کلی طول قدم‌ها می‌تواند متفاوت باشد اما حالتی را در نظر بگیرید که ذره در هر گام زمانی با احتمال p یک قدم با طول یک به سمت راست برمی‌دارد و با احتمال q یک قدم با طول یک به سمت چپ برمی‌دارد. بنابراین اگر قدم‌های این ذره را در گام زمانی iام با s_i نشان دهیم، بنابراین قدم‌های این ذره یک سری به شکل $s_1, s_2, \ldots$ هستند که نشان دهنده یک فرایند تصادفی است که در هر زمان مقدار آن ۱+ یا ۱− است. اگر موقعیت مکانی این ذره در زمان t_i با x_i نشان دهیم، می‌توانیم

[2] Markov process
[3] Hidden Markov models
[4] Stationary
[5] Random walk

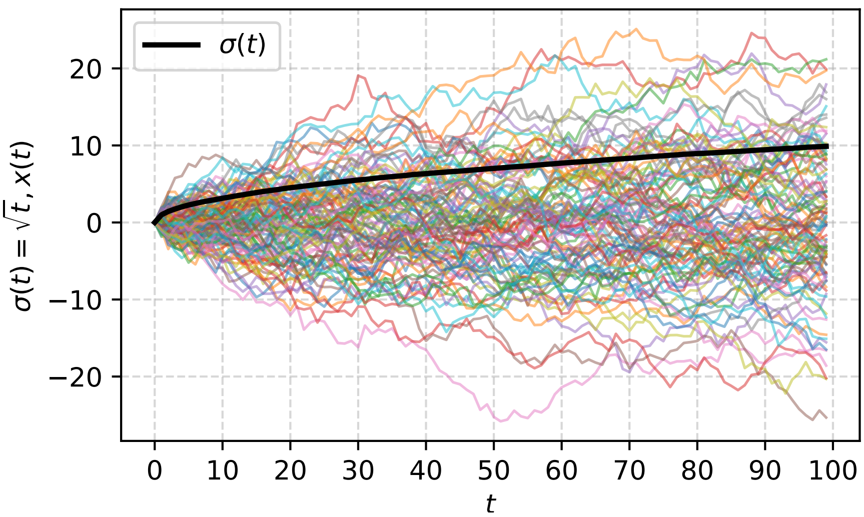
آن را به صورت زیر محاسبه کنیم.



شکل ۱.۲: ولگشت تصادفی ساده که با احتمال p قدمی به سمت راست و با احتمال q قدمی به سمت چپ برمی‌دارد.

$$x_n = \sum_{i=1}^{n} s_i$$

مسأله ولگشت تصادفی ساده را می‌توان در دو بعد و ابعاد بالاتر نیز بررسی کرد. در حالت کلی هیچ محدودیتی روی طول قدم‌ها وجود ندارد، مثلاً می‌توان حالتی را بررسی کرد طول قدم‌ها در هر زمان از یک توزیع گاوسی به شکل ۳۱.۱ پیروی کند. در این شرایط انحراف معیار توزیع مکان این ولگشت تصادفی به شکل $\sigma_x(t) = \sqrt{t}$ تغییر می‌کند که نشان دهنده مانا نبودن این فرایند تصادفی است. در شکل زیر مکان چند نمونه مستقل از ولگشت تصادفی با قدم‌های گاوسی کشیده شده است، منحنی سیاه رنگ نشان دهنده انحراف از معیار $\sigma_x(t)$ است.



شکل ۲.۲: چند نمونه مستقل از فرایند ولگشت تصادفی که قدم‌هایی با توزیع گاوسی برمی‌دارد.

۲.۲ فرایندهای مارکوف

همان‌طور که گفته شد در صورتی که توزیع احتمال توأم n نقطه‌ای یک فرایند تصادفی را داشته باشیم، همه اطلاعات ممکن برای یک فرایند تصادفی تا زمان t_n را می‌توانیم از روی توزیع احتمال توأم به دست آوریم. در حالتی که توزیع احتمال فرایند تصادفی در هر زمان به زمان‌های قبل از خودش وابستگی نداشته باشد و مستقل رفتار کند، توزیع احتمال توأم n نقطه‌ای از ضرب احتمال متغیر تصادفی در زمان‌های مختلف به دست می‌آید. یعنی:

$$p(x_n, t_n; x_{n-1}, t_{n-1} \dots ; x_0, t_0) = p(x_n, t_n)p(x_{n-1}, t_{n-1}) \dots p(x_0, t_0) \qquad (۳.۲)$$

به فرایندی که از رابطه بالا پیروی کند فرایند تصادفی مستقل گفته می‌شود. کاربرد این‌گونه فرایندها بسیار محدود است، بنابراین به مدل‌های تصادفی بهتری نیاز داریم. یکی از پرکاربرد ترین نوع فرایندهای تصادفی، فرایندهای مارکوف هستند. ایده اصلی فرایندهای مارکوف این است که حالت آینده یک فرایند تصادفی فقط به حالت کنونی فرایند وابسته است و از گام‌های زمانی قبل مستقل است پس می‌توان نتیجه گرفت که توزیع احتمال این نوع فرایندها در هر زمان فقط به زمان قبلی وابسته است، یعنی:

$$p(x_n, t_n \mid x_{n-1}, t_{n-1} \dots ; x_0, t_0) = p(x_n, t_n \mid x_{n-1}, t_{n-1}) \qquad (۴.۲)$$

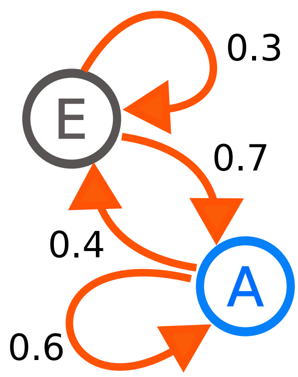
که رابطه بالا به ازای همه nها برقرار است. بنابراین می‌توان توزیع احتمال توأم n نقطه‌ای را به صورت زیر به دست آورد

$$\begin{aligned} &p(x_n, t_n; x_{n-1}, t_{n-1} \dots ; x_0, t_0) \\ &\quad = p(x_n, t_n \mid x_{n-1}, t_{n-1}) \dots p(x_2, t_2 \mid x_1, t_1)p(x_1, t_1 \mid x_0, t_0)p(x_0, t_0) \end{aligned} \qquad (۵.۲)$$

به احتمال‌های شرطی دو تایی در فرایندهای مارکوف، احتمال انتقال هم گفته می‌شود در شرایطی که این احتمال به زمان وابستگی نداشته باشد و فقط به اختلاف زمانی وابسته باشد به این فرایند، فرایند همگن[۶] گفته می‌شود. به عنوان مثالی از فرایندهای مارکوف همگن فرض کنید اگر امروز هوا آفتابی یا بارانی باشد بر روی شرایط آب و هوای فردا تأثیر گذار باشد و مهم نباشد که در چو موقعی از سال قرار داریم. شکل زیر نمایشی از ماتریس انتقال برای این

[6] Homogenous Markov chain

مثال است.



شکل ۳.۲: زنجیره مارکوف

اغلب به فرایندهای مارکوف با زمان گسسته و فضای نمونه شمارا، زنجیره مارکوف هم گفته می‌شود. در جدول زیر انواع فرایندهای مارکوف در شرایط مختلف آورده شده است، هرچند بر روی این تقسیم‌بندی و برخی اصطلاحات توافقی وجود ندارد. اما استفاده از نام زنجیره مارکوف برای فرایندهای با زمان و فضای نمونه گسسته رایج است. [۲۷]

جدول ۱.۲: تقسیم‌بندی فرایندهای مارکوف

	فضای نمونه شمارا	فضای نمونه پیوسته
زمان گسسته	زنجیره مارکوف	زنجیره مارکوف روی فضای نمونه قابل اندازه‌گیری
زمان پیوسته	فرایند مارکوف زمان‌پیوسته یا فرایند پرش مارکوف	دیگر فرایندهای تصادفی با خاصیت مارکوف

از این پس تمام روابط برای فرایند مارکوف با فضای نمونه و زمان پیوسته نوشته خواهند شد اما در ادامه به بررسی چند ویژگی مهم می‌پردازیم که بیشتر برای زنجیره‌های مارکوف استفاده می‌شوند.

۱.۲.۲ حد مانای فرایندهای مارکوف

برخی فرایندهای مارکوف پس از گذشت زمان کافی به حالتی مانا می‌رسند. در این بخش به بررسی روشی برای یافتن حالت مانا فرایندهای مارکوف می‌پردازیم که بیشتر برای زنجیره‌های مارکوف کاربردی است. با توجه به رابطه

۷.۱ می‌توانیم بنویسیم.

$$p(x_{i+1}, t_{i+1}; x_i, t_i) = p(x_{i+1}, t_{i+1} \mid x_i, t_i) p(x_i, t_i)$$

که x_i و x_{i+1}، متغیر تصادفی x را در دو زمان متوالی t_i و t_{i+1} نشان می‌دهد. با استفاده از رابطه ۱۶.۱ و جمع زدن بر روی x_i می‌توانیم توزیع احتمال x_{i+1} را در زمان t_{i+1} به دست آوریم.

$$p(x_{i+1}, t_{i+1}) = \int_{-\infty}^{+\infty} p(x_{i+1}, t_{i+1} \mid x_i, t_i) p(x_i, t_i) dx_i \tag{۶.۲}$$

رابطه بالا درواقع نشان دهنده ضرب ماتریس انتقال $\mathbf{p_{i+1,i}}$ در بردار توزیع احتمال $\mathbf{p_i}$ است.

$$\mathbf{p_{i+1}} = \mathbf{p_{i+1,i}} \times \mathbf{p_i}$$

که حروف پررنگ نشان‌دهنده شکل ماتریسی و برداری است. اگر $\mathbf{p_{i+2,i+1}}$ را در دو طرف رابطه بالا ضرب کنیم.

$$\mathbf{p_{i+2}} = \mathbf{p_{i+2,i+1}} \times \mathbf{p_{i+1,i}} \times \mathbf{p_i}$$

در حالتی که فرایند همگن باشد یعنی $p(x_{i+2}, t_{i+2} \mid x_{i+1}, t_{i+1}) = p(x_{i+1}, t_{i+1} \mid x_i, t_i)$ باشد و ماتریس انتقال را نیز با $\mathbf{W} = \mathbf{p_{i+2,i+1}} = \mathbf{p_{i+1,i}}$ نشان دهیم خواهیم داشت.

$$\mathbf{p_{i+2}} = \mathbf{W}^2 \times \mathbf{p_i}$$

رابطه بالا یکی از نتیاج مهم برای فرایندهای مارکوف همگن است، اگر این کار را ادامه دهیم می‌توانیم احتمال انتقال برای n گام زمانی را برای یک فرایند مارکوف همگن به دست آوریم. که به شکل زیر به دست می‌آید.

$$\mathbf{p_n} = \mathbf{W}^n \times \mathbf{p_0} \tag{۷.۲}$$

با فرض اینکه حالت مانای فرایند x وجود دارد و $p_s(x)$ نشان دهنده توزیع احتمال x در زمان t_s باشد یعنی زمانی که فرایند به حالت مانای خود رسیده است، اگر ماتریس انتقال را در این احتمال ضرب کنیم، توزیع به دست آمده باید برابر $p_s(x)$ باشد. [۲۸]

$$\mathbf{p_s}(\mathbf{x}) = \mathbf{W}^n \times \mathbf{p_s}(\mathbf{x})$$

رابطه بالا در فیزیک به معادله ویژه برداری معروف است، با حل این معادله می‌توان $p_s(x)$ را به دست آورد.

۲.۲.۲ معادله چَپمَن-کولموگروف

با توجه به رابطه ۱.۱۶ دیدیم که توزیع‌های احتمال توأم چگونه به هم مرتبط هستند اگر بخواهیم احتمال توأم دونقطه‌ای را از احتمال توأم سه نقطه‌ای به دست آوریم داریم:

$$p\left(x_3, t_3; x_1, t_1\right) = \int_{-\infty}^{+\infty} p\left(x_3, t_3; x_2, t_2; x_1, t_1\right) dx_2$$

اگر x یک فرایند مارکوف باشد بنابراین $p(x_3, t_3; x_2, t_2; x_1, t_1) = p(x_3, t_3 \mid x_2, t_2) p(x_2, t_2 \mid x_1, t_1) p(x_1, t_1)$ با جایگذاری در رابطه بالا داریم:

$$p\left(x_3, t_3; x_1, t_1\right) = \int_{-\infty}^{+\infty} p\left(x_3, t_3 \mid x_2, t_2\right) p\left(x_2, t_2 \mid x_1, t_1\right) p\left(x_1, t_1\right) dx_2 \tag{۲.۸}$$

معادله بالا به معادله چپمن-کولموگروف [۷] معروف است و باید توجه شود که معادله بالا برای سه زمان متوالی t_1، t_2 و t_3 نوشته شده است. اگر دو طرف تساوی را به $p(x_1, t_1)$ تقسیم کنیم با توجه به رابطه احتمال شرطی و توأم ۱.۶ به شکل دیگری از معادله چپمن-کولموگروف می‌رسیم.

$$p\left(x_3, t_3 \mid x_1, t_1\right) = \int_{-\infty}^{+\infty} p\left(x_3, t_3 \mid x_2, t_2\right) p\left(x_2, t_2 \mid x_1, t_1\right) dx_2 \tag{۲.۹}$$

معادله چپمن-کولموگروف برای همه‌ی فرایندهای مارکوف برقرار است، اما هر فرایندی که معادله چپمن-کولموگروف برای آن برقرار باشد لزوماً مارکوف نیست. به عبارت دیگر برقرار بودن معادله چپمن-کولموگروف شرط لازم برای

[7] Chapman-Kolmogorov equation

یک فرایند مارکوف است.

۳.۲.۲ فرایند وینر

یک مثال پرکاربرد از فرایندهای مارکوف، فرایند وینر [۸] است. فرایند وینر از بسیاری جهات مشابه یک ولگشت تصادفی با قدم‌های گاوسی است، با این تفاوت که توزیع قدم در هر زمان به قدم قبلی وابسته است. احتمال انتقال فرایند وینر به شکل زیر است:

$$p_2\left(x, t \mid x', t'\right) = \frac{1}{\sqrt{2\pi\tau}} e^{-\left(x-x'\right)^2/2\tau} \tag{۱۰.۲}$$

که $\tau = t - t'$ است، بنابراین فرایند وینر یک فرایند مارکوف همگن است که احتمال انتقال آن فقط به τ وابسته است. توزیع احتمال این فرایند در هر زمان به شکل زیر است.

$$p_1(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \tag{۱۱.۲}$$

فرایند وینر را در فیزیک می‌توان به عنوان فرایند پخش گرما یا حرکت براونی ذرات تعبیر کرد. نکته مهم در مورد فرایند وینر این است که این فرایند حد مانا ندارد. [۱۸]

۴.۲.۲ معادله مادر

برای فهمیدن رفتار یک فرایند مارکوف معادله چپمن-کولموگروف کمک زیادی نخواهد کرد اما می‌توان از معادله چپمن-کولموگروف به معادله‌ای سودمند به نام معادله مادر [۹] رسید. معادله مادر یک معادله دیفرانسیلی برای احتمال انتقال است که تغییرات احتمال انتقال بر حسب زمان را به دست می‌دهد. برای به دست آوردن این معادله باید رفتار احتمال انتقال را در اختلاف زمان‌های کوتاه یعنی $\Delta t \to 0$ مشخص کنیم. در معادله ۹.۲ به ازای $t_2 = t_1 = t$ به

[8] Wiener process

[9] Master equation

نتیجه بدیهی زیر می‌رسیم.

$$\begin{aligned} p\left(x_3, t_3 \mid x_1, t_1\right) &= \int_{-\infty}^{+\infty} p\left(x_3, t_3 \mid x_2, t\right) p\left(x_2, t \mid x_1, t\right) dx_2 \\ &\Rightarrow p\left(x_2, t \mid x_1, t\right) = \delta\left(x_2 - x_1\right) \end{aligned}$$

که تابع دلتا به دست آمده درواقع مرتبه صفرم از رفتار $p(x', t' \mid x, t)$ در زمان‌های کوتاه است. با در نظر داشتن این نکته می‌توان رابطه زیر را برای رفتار احتمال انتقال در زمان‌های کوتاه به دست آورد.

$$p\left(x_2, t + \Delta t \mid x_1, t\right) = \delta\left(x_2 - x_1\right)\left[1 - a^{(0)}\left(x_1, t\right) \Delta t\right] + w_t\left(x_2 \mid x_1\right) \Delta t + O\left[(\Delta t)^2\right] \quad (۱۲.۲)$$

در رابطه بالا $w_t(x_2 \mid x_1)$ را به عنوان احتمال انتقال از x_1 به x_2 بر واحد زمان t تعبیر می‌کنیم. بنابراین ضریب $1 - a^{(0)}\left(x_1, t\right) \Delta t$ را باید احتمال اینکه هیچ انتقالی در مدت زمان Δt صورت نگیرد تعبیر کرد یعنی احتمال اینکه در این مدت زمان x ثابت باشد. از نرمال بودن احتمال انتقال $p(x_2, t_2 \mid x_1, t_1)$ می‌توانیم بنویسیم.

$$1 = \int \mathrm{d}x_2 P\left(x_2, t + \Delta t \mid x_1, t\right) \simeq 1 - a^{(0)}\left(x_1, t\right) \Delta t + \int \mathrm{d}x_2 w_t\left(x_2 \mid x_1\right) \Delta t$$

بنابراین داریم

$$a^{(0)}\left(x_1, t\right) = \int \mathrm{d}x_2 w_t\left(x_2 \mid x_1\right) \quad (۱۳.۲)$$

از رابطه بالا می‌بینیم که $a^{(0)}\left(x_1, t\right) \Delta t$ احتمال کل انتقال از x_1 در بازه $(t, t + \Delta t)$را نشان می‌دهد بنابراین $1 - a^{(0)}\left(x_1, t\right) \Delta t$ احتمال این است که هیچ انتقالی در این بازه اتفاق نیفتد.

حالا می‌توانیم با استفاده از معادله چپمن-کولموگروف ۹.۲ معادله دیفرانسیلی برای احتمال انتقال به دست آوریم. با جایگذاری رابطه ۱۲.۲ در معادله چپمن-کولموگروف ۹.۲ به ازای $t_3 = t_2 + \Delta t$ داریم:

$$p(x_3,t_2+\Delta t \mid x_1,t_1)=\int \mathrm{d}x_2\, \overbrace{p(x_3,t_2+\Delta t \mid x_2,t_2)}^{\delta(x_3-x_2)\left[1-a^{(0)}(x_2,t_2)\Delta t\right]+w_{t_2}(x_3|x_2)\Delta t}\, p(x_2,t_2 \mid x_1,t_1)$$
$$\simeq \left[1-a^{(0)}(x_3,t_2)\,\Delta t\right] p(x_3,t_2 \mid x_1,t_1)$$
$$+\,\Delta t \int \mathrm{d}x_2 w_{t_2}(x_3 \mid x_2)\, p(x_2,t_2 \mid x_1,t_1)$$

سپس $a^{(0)}(x_3,t_2)$ را با استفاده از ۱۳.۲ بر حسب $w_{t_2}(x_2 \mid x_3)$ می‌نویسیم. بنابراین رابطه بالا را به شکل زیر می‌نویسیم.

$$\frac{1}{\Delta t}\left[p(x_3,t_2+\Delta t \mid x_1,t_1)-p(x_3,t_2 \mid x_1,t_1)\right]$$
$$\simeq \int \mathrm{d}x_2\left[w_{t_2}(x_3 \mid x_2)\, p(x_2,t_2 \mid x_1,t_1)-w_{t_2}(x_2 \mid x_3)\, p(x_3,t_2 \mid x_1,t_1)\right]$$

نشان گذاری را با تبدیل x_1 و t_1 به x_0 و t_0، x_2 و t_2 را به x' و t' و در نهایت x_3 را به x تغییر می‌دهیم. بنابراین در حد $\Delta t\to 0$ می‌توانیم معادله مادر را به شکل زیر به دست آوریم.

$$\frac{\partial}{\partial t}p(x,t \mid x_0,t_0)=\int \mathrm{d}x'\left[w_t(x|x')\, p(x',t \mid x_0,t_0)-w_t(x' \mid x)\, p(x,t \mid x_0,t_0)\right] \quad (۱۴.۲)$$

که یک معادله انتگرال - دیفرانسیل است. معادلات انتگرال - دیفرانسیل معادلاتی هستند که شامل انتگرال و مشتق یک تابع هستند.

معادله مادر یک شکل دیفرانسیلی از معادله چپمن–کولموگروف است. بنابراین نمی‌توان با آن توزیع احتمال $p(x,t)$ را حساب کرد و برای محاسبه احتمال انتقال $p(x,t \mid x_0,t_0)$ کاربرد دارد. اما می‌توان نشان داد که با در نظر گرفتن شرایط اولیه می‌توان معادله‌ای برای توزیع احتمال با استفاده از معادله مادر به شکل زیر به دست آورد. [۲۶]

$$\frac{\partial p(x,t)}{\partial t}=\int \mathrm{d}x'\left[w(x \mid x')\, p(x',t)-w(x' \mid x)\, p(x,t)\right] \quad (۱۵.۲)$$

از شکل معادله مادر می‌بینیم که این معادله درواقع یک معادله توازن برای احتمال حالت‌های (مقادیر) ممکن x است.

۵.۲.۲ بسط کرامرز-مویال و معادله فوکر-پلانک

بسط کرامرز-مویال[۱۰] معادله مادر را از یک معادله انتگرال - دیفرانسیل به یک معادله دیفرانسیل با مرتبه بی‌نهایت تبدیل می‌کند. بنابراین استفاده از این بسط از معادله مادر آسان‌تر نیست اما تحت شرایطی ممکن است که بتوان با تعداد محدودی از جملات بسط معادله را حل کرد. [۲۹]

در ابتدا با یک تغییر متغیر احتمال انتقال w را به تابعی بر حسب r که $r = x' - x$ تبدیل می‌کنیم. یعنی احتمال انتقال تابعی است از میزان جابه‌جایی r و نقطه شروع x'، بنابراین داریم:

$$w\left(x \mid x'\right) = w\left(x'; r\right), \quad r = x - x' \tag{۱۶.۲}$$

بنابراین معادله مادر ۱۵.۲ را می‌توان به شکل زیر نوشت.

$$\frac{\partial p(x,t)}{\partial t} = \int w(x-r;r)p(x-r,t)\mathrm{d}r - p(x,t)\int w(x;-r)\mathrm{d}r \tag{۱۷.۲}$$

که تغییر علامت مربوط به تغییر متغیر $x' \to r = x - x'$ در کران‌های انتگرال به این شکل تأثیر گذاشته است که

$$\int_{-\infty}^{\infty}\mathrm{d}x' f\left(x'\right) = -\int_{x+\infty}^{x-\infty}\mathrm{d}r f(x-r) = -\int_{\infty}^{-\infty}\mathrm{d}r f(x-r) = \int_{-\infty}^{\infty}\mathrm{d}r f(x-r)$$

با فرض اینکه تغییرات x به شکل پرش‌های کوچک اتفاق می‌افتد بنابراین $w(x';r)$ یک تابع تیز از r است اما تغییرات آن با x' به اندازه کافی آرام است. فرض دیگری که می‌کنیم این است که $p(x,t)$ نیز به آرامی با x تغییر می‌کند. حال می‌توانیم تغییر x به $x-r$ را با بسط تیلور بررسی کنیم بنابراین جمله اول سمت راست رابطه ۱۷.۲ را بسط تیلور می‌دهیم.

$$\begin{aligned}
\frac{\partial p(x,t)}{\partial t} &= \int w(x;r)p(x,t)\mathrm{d}r + \sum_{m=1}^{\infty}\frac{(-1)^m}{m!}\int r^m\frac{\partial^m}{\partial x^m}[w(x;r)p(x,t)]\mathrm{d}r \\
&\quad - p(x,t)\int w(x;-r)\mathrm{d}r \\
&= \sum_{m=1}^{\infty}\frac{(-1)^m}{m!}\frac{\partial^m}{\partial x^m}\left\{\left[\int r^m w(x;r)\mathrm{d}r\right]p(x,t)\right\}
\end{aligned}$$

[10] Kramers–Moyal expansion

که جمله اول و سوم در رابطه بالا با یکدیگر ساده شده‌اند بنابراین داریم. در انتها ضرایب بسط به شکل زیر تعریف می‌کنیم.

$$a^{(m)}(x,t) = \int r^m w(x;r)\mathrm{d}r \tag{۱۸.۲}$$

بنابراین بسط کرامرز–مویال معادله مادر به شکل زیر به دست می‌آید.

$$\frac{\partial p(x,t)}{\partial t} = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \frac{\partial^m}{\partial x^m} \left[a^{(m)}(x,t)p(x,t)\right] \tag{۱۹.۲}$$

بسط کرامرز–مویال هیچ تفاوتی با معادله مادر ندارد پس استفاده از این بسط به اندازه معادله مادر مشکل است. اما همان‌طور که گفته شد ممکن است تحت شرایطی بتوان از جملات مرتبه بالا بسط کرامز مویال صرف نظر کرد. مثلاً در شرایطی که به ازای $m > 2$ ضرایب $a^m(x,t)$ خیلی کوچک باشند و بتوان از آن‌ها صرف نظر کرد، بسط کرامرز–مویال به شکل زیر تبدیل می‌شود.

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x} \left[a^{(1)}(x,t)p(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2} \left[a^{(2)}(x,t)p(x,t)\right] \tag{۲۰.۲}$$

به معادله بالا، معادله فوکر–پلانک[11] گفته می‌شود. به جمله اول این معادله جمله رانش و به جمله دوم آن جمله پخش گفته می‌شود، همچنین به ضرایب $a^1(x,t)$ و $a^2(x,t)$ ضریب رانش و ضریب پخش گفته می‌شود. یادآوری این نکته لازم است که معادله فوکر–پلانک به عنوان حالت خاصی از بسط کرامرز–مویال از معادله مادر به دست می‌آید، بنابراین این معادله نیز رفتار احتمال انتقال را در زمان نشان می‌دهد اما همان‌طور که گفته شد با در نظر گرفتن شرایط اولیه می‌توان رفتار توزیع احتمال را همان‌طور که در ۲۰.۲ نوشته شده است نیز به دست آورد. [۳۰]

۶.۲.۲ ممان‌های شرطی

احتمال انتقال بر واحد زمان $w(x' \mid x)$ در تعریف ممان‌های شرطی ۱۸.۲ استفاده شده است. بنابراین مجبوریم برای محاسبه $a^m(x.t)$ از رابطه ۱۲.۲ که رابطه بین $w(x' \mid x)$ و احتمال انتقال برای اختلاف زمان‌های کوچک

[11] Fokker–Planck equation

است، استفاده کنیم. در رابطه ۱۶.۲ دیدیم که $w(x';r) = w(x \mid x')$ و $x = x' + r$ بر همین اساس می‌توانیم بنویسیم:

$$w(x;r) = w\left(x' \mid x\right), \qquad x' = x + r$$

با جایگذاری این عبارت در معادله ۱۸.۲ می‌توان نوشت.

$$a^{(m)}(x,t) = \int \mathrm{d}x' \left(x' - x\right)^m w\left(x' \mid x\right) \tag{۲۱.۲}$$

به منظور اینکه رابطه‌ای برای ضرایب به دست بیاوریم کمیت زیر را معرفی می‌کنیم.

$$\mathcal{A}^{(m)}(x;\tau,t) = \int \mathrm{d}x' \left(x' - x\right)^m p\left(x', t+\tau \mid x, t\right) \tag{۲۲.۲}$$

که درواقع میانگین $[X(t+\tau) - X(t)]^m$ به ازای $X(t) = x$ است و به آن ممان شرطی می‌گویند. سپس با استفاده از احتمال انتقال برای زمان‌های کوتاه ۱۲.۲ می‌توانیم بنویسیم

$$\begin{aligned}
\mathcal{A}^{(m)}(x;\tau,t) &= \int \mathrm{d}x' \left(x' - x\right)^m \left\{\delta\left(x' - x\right)\left[1 - a^{(0)}(x,t)\tau\right] + w\left(x' \mid x\right)\tau + \mathcal{O}\left(\tau^2\right)\right\} \\
&= \tau \int \mathrm{d}x' \left(x' - x\right)^m w\left(x' \mid x\right) + \mathcal{O}\left(\tau^2\right) \\
&= a^{(m)}(x,t)\tau + \mathcal{O}\left(\tau^2\right), \qquad (m \geq 1)
\end{aligned}$$

که اولین جمله انتگرال به دلیل وجود تابع دلتای دیراک از بین می‌رود. بنابراین می‌توان ضرایب بسط را از مشتق ممان‌های شرطی به شکل زیر به دست آورد

$$a^{(m)}(x,t) = \frac{\partial}{\partial \tau} \mathcal{A}^{(m)}(x;\tau,t)\bigg|_{\tau=0} \tag{۲۳.۲}$$

در انتها با نوشتن $\mathcal{A}$ به شکل زیر:

$$\mathcal{A}^{(m)}(x;\Delta t,t) = \int \mathrm{d}x' \left(x' - x\right)^m p\left(x', t+\Delta t \mid x, t\right) = \langle [X(t+\Delta t) - X(t)]^m \rangle|_{X(t)=x}$$

می‌توانیم ضرایب بسط را با استفاده از رابطه زیر به دست آوریم.

$$a^{(m)}(x,t) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\langle [X(t+\Delta t)-X(t)]^m\right\rangle\Big|_{X(t)=x} \tag{۲۴.۲}$$

۷.۲.۲ قضیه پاوولا

همان‌طور که دیدیم بسط کرامرز-مویال دارای تعداد نامتناهی جمله است که در حالت کلی حل این بسط را نسبت به معادله مادر ساده‌تر نمی‌کند، اما اگر بتوان تعدادی از جملات را نگه داشت و از بقیه آن‌ها صرف نظر کرد حل بسط کرامرز-مویال بسیار ساده‌تر از معادله مادر خواهد بود. برای اینکه بتوانیم از این جملات صرف نظر کنیم، ممان‌های متعلق به این جملات باید صفر باشند. اما بررسی کردن تعداد نامتناهی ممان به اندازه حل بسط کرامرز-مویال کار دشواری است. با روشی که قضیه پاوولا[12] معرفی می‌کند می‌توانیم از صفر بودن ممان‌ها، بدون بررسی تک‌تک آن‌ها مطمئن شویم. بدین منظور باید از حالت کلی نامساوی شوارتز استفاده کنیم که به شکل زیر است.

$$[f(x)g(x)P(x)\mathrm{d}x]^2 \le \int f^2(x)P(x)\mathrm{d}x\int g^2(x)P(x)\mathrm{d}x \tag{۲۵.۲}$$

در نامساوی بالا $P(x)$ یک تابع مثبت است و توابع $f(x)$ و $g(x)$ توابع دلخواهی هستند.

$$\begin{aligned} &f(x) = \left(x-x'\right)^n; \qquad g(x) = \left(x-x'\right)^{n+m}\\ &P(x) = P\left(x,t+\tau|x',t'\right)\end{aligned}$$

به ازای توابع به شکل بالا نامساوی شوارتز به عبارت زیر که بین ممان‌های شرطی است تبدیل می‌شود. [۱۸]

$$\mathcal{A}_{2n+m}^2 \le \mathcal{A}_{2n}\cdot\mathcal{A}_{2n+2m} \tag{۲۶.۲}$$

برای $n=0$ داریم $\mathcal{A}_m \le \mathcal{A}_{2m}$ که این رابطه به وضوح برای $m=0$ برقرار است. از این رابطه هیچ محدودیتی برای ضرایب بسط $a^{(m)}(x,t)$ به دست نمی‌آید. به ازای $m=0$ رابطه ۲۶.۲ به $\mathcal{A}_{2n}^2 \le \mathcal{A}_{2n}^2$ تبدیل می‌شود که به ازای همه nها برقرار است. بنابراین ناچاریم که رابطه ۲۶.۲ را به ازای $n\ge 1$ و $m\ge 1$ بررسی کنیم. از دو طرف

[12] Pawula theorem

رابطه ۲۶.۲ نسبت به τ مشتق می‌گیریم و به ازای $\tau = 0$ این رابطه به رابطه‌ای مشابه برای ضرایب بسط $a^{(m)}(x,t)$ تبدیل می‌شود.

$$[a^{(2n+m)}]^2 \leq a^{(2n)} \cdot a^{(2n+2m)} \tag{۲۷.۲}$$

اگر $a^{(2n)} = 0$ صفر باشد بنابراین $a^{(2n+m)}$ هم حتماً صفر است، به عبارت دیگر:

$$a^{(2n)} = 0 \Rightarrow a^{(2n+1)} = a^{(2n+2)} = \ldots = 0 \quad (n \geq 1) \tag{۲۸.۲}$$

علاوه بر این اگر $a^{(2n+2m)}$ صفر باشد، $a^{(2n+m)}$ هم حتماً صفر است، به عبارت دیگر اگر $r = n + m$ در نظر بگیریم، داریم:

$$\begin{aligned} a^{(2r)} = 0 &\Rightarrow a^{(r+n)} = 0 \quad (n = 1, \ldots, r-1) \\ &\Rightarrow a^{(2r-1)} = a^{(2r-2)} = \ldots = a^{(r+1)} = 0 \quad (r \geq 2) \end{aligned} \tag{۲۹.۲}$$

با استفاده از ۲۸.۲ و ۲۹.۲ می‌توان به این نتیجه رسید که به ازای $r \geq 1$ اگر داشته باشیم $a^{(2r)} = 0$ تمام ضرایب $a^{(n)}$ به ازای $n \geq 3$ نیز صفر خواهند شد. به عبارت دیگر:

$$a^{(2n)} = 0 \Rightarrow a^{(3)} = a^{(4)} = \ldots = 0 \quad (n \geq 1) \tag{۳۰.۲}$$

رابطه بالا چیزی است که به آن قضیه پاوولا گفته می‌شود و اگر شرط بالا برقرار باشد بسط کرامرز-مویال به معادله فوکر-پلانک تبدیل می‌شود. [۳۱]

۳.۲ معادله لانژوین

تا اینجا فرایندهای تصادفی را با استفاده از توزیع احتمال آن‌ها بررسی کرده‌ایم. در این بخش روش متفاوتی را در پیش می‌گیریم و به جای بررسی تحول زمانی توزیع احتمال، به شکل مستقیم تحول زمانی متغیر تصادفی را بررسی

می‌کنیم. رابطه زیر شکل کلی معادله لانژوین[۱۳] است. [۳۲]

$$\frac{\mathrm{d}x}{\mathrm{d}t} = A(x,t) + B(x,t)\zeta(t) \tag{۳۱.۲}$$

در رابطه بالا $\zeta(t)$ یک فرایند تصادفی دلخواه است که به آن نیروی تصادفی هم گفته می‌شود. انتخاب $\zeta(t)$ مناسب باعث می‌شود که $x(t)$ یک فرایند مارکوف باشد. اگر $\zeta(t)$ توزیع گاوسی با آمار زیر باشد، متغیر تصادفی $x(t)$ خاصیت مارکوف خواهد داشت.

$$\begin{aligned} \langle \zeta(t) \rangle &= 0 \\ \langle \zeta(t_1)\, \zeta(t_2) \rangle &= 2D\delta\,(t_1 - t_2) \end{aligned} \tag{۳۲.۲}$$

از آنجایی که معادله ۳۱.۲ یک معادله دیفرانسیل مرتبه اول است، هر نمونه از $\zeta(t)$ با مشخص بودن شرایط اولیه $x(t_0)$، $x(t)$ به صورت یکتا مشخص می‌شود. علاوه بر این ، مقادیر جمله دوم در زمان‌های متفاوت مستقل از هم هستند، به دلیل اینکه همبستگی $\zeta(t)$ به شکل تابع دلتای دیراک است. بنابراین مقادیر $\zeta(t)$ در زمان‌های قبل مثلاً $t' < t_0$ نمی‌توانند تأثیری بر احتمال‌های شرطی در زمان‌های $t > t_0$ بگذارند. با توجه به شرایط گفته شده می‌توان نتیجه گرفت که جواب معادله لانژوین ۳۱.۲ دارای خاصیت مارکوف است.

جمله $A(x,t)$ و $B(x,t)\zeta(t)$ را اغلب با نام جمله رانش و جمله پخش می‌شناسند. به دلیل وجود $\zeta(t)$ معادله ۳۱.۲ یک معادله دیفرانسیل تصادفی است که به معنی وجود جملات تصادفی در معادله است که باعث به وجود آمدن خواص تصادفی در $x(t)$ می‌شود. حل یک معادله لانژوین به معنی یافتن خواص آماری فرایند $x(t)$ است.

۱.۳.۲ ضرایب بسط کرامرز-مویال و معادله لانژوین

از آنجایی که جواب معادله لانژوین یک فرایند مارکوف است، بنابراین باید از یک معادله مادر پیروی کند که می‌توانیم آن را به شکل بسط کرامرز-مویال ۱۹.۲ بنویسیم. حال می‌خواهیم ضرایب بسط ۲۴.۲ را برای یک معادله لانژوین

[13] Langevin equation

به دست آوریم. برای این کار معادله ۳۱.۲ را به یک معادله انتگرالی تبدیل می‌کنیم.

$$x(t+\Delta t)-x=\int_t^{t+\Delta t}\mathrm{d}t_1 A\left[x\left(t_1\right),t_1\right]+\int_t^{t+\Delta t}\mathrm{d}t_1 B\left[x\left(t_1\right),t_1\right]\zeta\left(t_1\right) \tag{۳۳.۲}$$

که x به معنی مقدار اولیه $x(t)$ است، ضرایب رانش و پخش معادله لانژوین را به شکل زیر بسط می‌دهیم.

$$\begin{aligned}
A\left[x\left(t_1\right),t_1\right]&=A\left(x,t_1\right)+A'\left(x,t_1\right)\left[x\left(t_1\right)-x\right]+\cdots\\
B\left[x\left(t_1\right),t_1\right]&=B\left(x,t_1\right)+B'\left(x,t_1\right)\left[x\left(t_1\right)-x\right]+\cdots
\end{aligned}$$

ضرایب با علامت پرایم درواقع مشتقات جزئی نسبت به در نقطه اولیه x هستند.

$$A'(x,t)\equiv\left.\frac{\partial A}{\partial x}\right|_x \qquad B'(x,t)\equiv\left.\frac{\partial B}{\partial x}\right|_x$$

بنابراین می‌توانیم بنویسیم

$$\begin{aligned}
x(t+\Delta t)-x=&\int_t^{t+\Delta t}\mathrm{d}t_1 A\left(x,t_1\right)\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 A'\left(x,t_1\right)\left[x\left(t_1\right)-x\right]+\cdots\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 B\left(x,t_1\right)\zeta\left(t_1\right)\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 B'\left(x,t_1\right)\left[x\left(t_1\right)-x\right]\zeta\left(t_1\right)+\cdots
\end{aligned}$$

برای محاسبه $x(t_1)-x$ دوباره از رابطه ۳۳.۲ در رابطه بالا استفاده می‌کنیم.

$$\begin{aligned}
x(t+\Delta t)-x=&\int_t^{t+\Delta t}\mathrm{d}t_1 A\left(x,t_1\right)+\int_t^{t+\Delta t}\mathrm{d}t_1 A'\left(x,t_1\right)\int_t^{t_1}\mathrm{d}t_2 A\left(x,t_2\right)\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 A'\left(x,t_1\right)\int_t^{t_1}\mathrm{d}t_2 A\left(x,t_2\right)\zeta\left(t_2\right)+\cdots\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 B\left(x,t_1\right)\zeta\left(t_1\right)+\int_t^{t+\Delta t}\mathrm{d}t_1 B'\left(x,t_1\right)\zeta\left(t_1\right)\int_t^{t_1}\mathrm{d}t_2 A\left(x,t_2\right)\\
&+\int_t^{t+\Delta t}\mathrm{d}t_1 B'\left(x,t_1\right)\zeta\left(t_1\right)\int_t^{t_1}\mathrm{d}t_2 B\left(x,t_2\right)\zeta\left(t_2\right)+\cdots
\end{aligned}$$

اگر مقدار چشم‌داشتی رابطه بالا را به ازای $x(t) = x$، با استفاده از خواص آماری ۳۲.۲ حساب کنیم، میانگین شرطی که برای محاسبه $a^{(1)}(x,t)$ نیاز است به دست می‌آید.

$$\begin{aligned}\langle x(t+\Delta t)-x\rangle = &\int_t^{t+\Delta t} \mathrm{d}t_1 A\left(x,t_1\right)+\int_t^{t+\Delta t} \mathrm{d}t_1 A'\left(x,t_1\right)\int_t^{t_1} \mathrm{d}t_2 A\left(x,t_2\right)\\ &+2D\int_t^{t+\Delta t} \mathrm{d}t_1 B'\left(x,t_1\right)\int_t^{t_1} \mathrm{d}t_2 B\left(x,t_2\right)\delta\left(t_2-t_1\right)+\cdots\end{aligned}$$

سپس با توجه به رابطه $\int_{t_0}^{t_1} \mathrm{d}t\delta\left(t-t_0\right)f(t)=\frac{1}{2}f\left(t_0\right)$، می‌توانیم بنویسیم.

$$\int_t^{t_1} \mathrm{d}t_2 B\left(x,t_2\right)\delta\left(t_2-t_1\right)=\frac{1}{2}B\left(x,t_1\right) \tag{۳۴.۲}$$

در انتها با توجه به اینکه $a^{(1)} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle x(t+\Delta t)-x\rangle\Big|_{x(t)=x}$ و از آنجایی که تنها به جملات با مرتبه Δt نیاز داریم می‌توانیم بنویسیم.

$$a^{(1)}(x,t)=A(x,t)+DB(x,t)\frac{\partial B(x,t)}{\partial x}$$

دیگر انتگرال‌هایی که در رابطه بالا نوشته نشده‌اند در حد $\Delta \to 0$ قابل چشم‌پوشی هستند.

با استفاده از روشی مشابه می‌توانیم دومین ضریب بسط کرامرز–مویال $a^{(2)}(x,t)$ را به دست آوریم.

$$a^{(2)}(x,t)=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_t^{t+\Delta t} \mathrm{d}t_1 B\left(x,t_1\right)\int_t^{t+\Delta t} \mathrm{d}t_2 B\left(x,t_2\right)2D\delta\left(t_1-t_2\right)=2DB^2(x,t)$$

در حالی که تمام ضرایب $a^{(m)}$ برای $m \geq 3$ صفر می‌شوند. بنابراین با توجه به نتایج به دست آمده می‌توانیم بنویسیم.

$$\begin{aligned}&a^{(1)}(x,t)=A(x,t)+DB(x,t)\frac{\partial B(x,t)}{\partial x}\\&a^{(2)}(x,t)=2DB^2(x,t)\\&a^{(m)}(x,t)=0, \quad \text{for} \quad m\geq 3\end{aligned} \tag{۳۵.۲}$$

۲.۳.۲ معادله فوکر-پلانک و معادله لانژوین

با توجه به رابطه ۳۵.۲ می‌توان نتیجه گرفت که برای یک فرایند مارکوف که با استفاده از معادله لانژوین ۳۱.۲ تعیین شده است و همبستگی جمله تصادفی آن به شکل دلتای دیراک است، بسط کرامرز-مویال تا جملات مرتبه ۲ را شامل می‌شود. بنابراین توزیع احتمال این فرایند از معادله فوکر-پلانک ۲۰.۲ پیروی می‌کند، معادله فوکر-پلانک را می‌توان به شکل زیر نوشت.

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\left\{\left[A(x,t) + DB(x,t)\frac{\partial B(x,t)}{\partial x}\right]P\right\} + D\frac{\partial^2}{\partial x^2}\left[B^2(x,t)P\right] \tag{۳۶.۲}$$

ذکر این نکته لازم است که در کنار جمله رانش $A(x,t)$، $a^{(1)}(x,t)$ شامل یک جمله به شکل $DB(x,t)B'(x,t)$ نیز هست که به آن جمله رانش ناشی از نوفه هم گفته می‌شود. این معادله از این جهت حائز اهمیت است که این امکان را فراهم می‌کند تا معادله فوکر-پلانک را به شکل مستقیم با استفاده از ضرایبی که به طور مستقیم در معادله حرکت استفاده می‌شوند به دست آورد. [۲۶]

۴.۲ فرایندهای غیرمارکوف

همانطور که در بخش ۲.۲ دیدیم، فرایندهای مارکوف به فرایندهایی گفته می‌شود که رابطه ۴.۲ یا معادل آن رابطه ۵۸.۲ برای آن‌ها برقرار باشد و اگر این رابطه برای فرایندی برقرار نباشد به آن غیرمارکوف گفته می‌شود. به طور خاص فرایندهایی به شکل زیر را در نظر بگیرید.

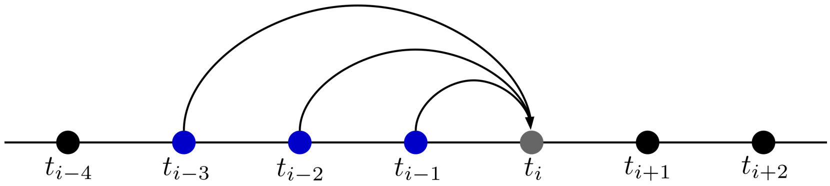
$$\begin{aligned} p(x_n, t_n; x_{n-1}, t_{n-1} \dots ; x_0, t_0) =& p(x_n, t_n \mid x_{n-1}, t_{n-1}; \dots x_0, t_0) \\ & \cdots p(x_1, t_1 \mid x_0, t_0) p(x_0, t_0) \end{aligned} \tag{۳۷.۲}$$

در واقع چنین فرایندی به تمامی حالت‌های گذشته خودش تا حالت اولیه وابسته است. ممکن است فرایندهایی وجود داشته باشند که تنها به تعداد متناهی از حالت‌های گذشته خود وابسته هستند؛ برای این‌گونه از فرایندها رابطه

زیر برقرار است.

$$p(x_n, t_n \mid x_{n-1}, t_{n-1} \dots ; x_0, t_0) = p(x_n, t_n \mid x_{n-1}, t_{n-1}; \dots x_{n-l}, t_{n-l}) \qquad (۳۸.۲)$$

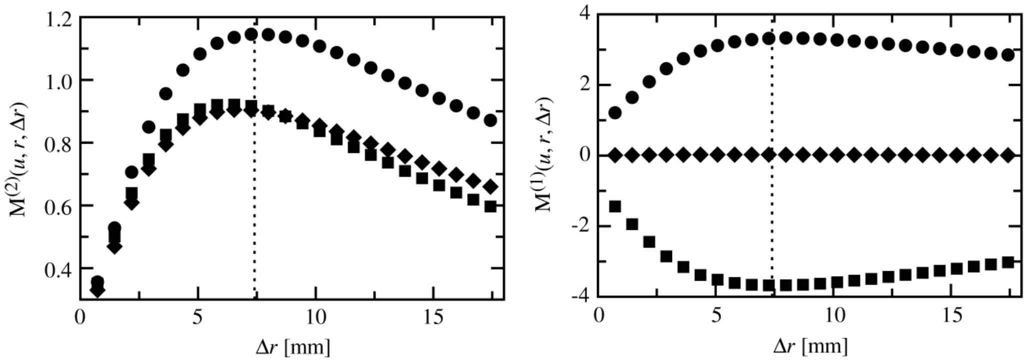
فرایندهایی که رابطه بالا برای آن‌ها برقرار باشد فرایندهای غیرمارکوفی هستند که اصطلاحاً حافظه‌ای به اندازه l زمان قبل دارند، به این نوع فرایندها، فرایندهای مارکوف مرتبه l نیز گفته می‌شود.



شکل ۴.۲: یک فرایند غیر مارکوف که حافظه‌ای به اندازه سه گام قبلی خود دارد

۱.۴.۲ طول مارکوف

اصولا تشخیص مارکوف بودن یا نبودن یک فرایند از روی داده‌های واقعی کار مشکلی است و دلیل اصلی آن این است که معمولا خاصیت مارکوف در فاصله‌های زمانی کوچک برقرار نیست. طول مارکوف به کوچک‌ترین مقیاسی گفته می‌شود که می‌توان یک فرایند تصادفی را مارکوف در نظر گرفت. یک فرایند تصادفی می‌تواند طول مارکوف متناهی یا نامتناهی داشته باشد. [۱۹] اگر بتوانیم چنین طولی را برای یک فرایند تصادفی پیدا کنیم می‌توانیم با استفاده از معادلات به دست آمده برای فرایندهای مارکوف دینامیک سیستم را با توجه به طول مارکوف بررسی کنیم. به عنوان مثال با توجه به رابطه ممان‌های شرطی با ضرایب بسط کرامرز-مویال یعنی رابطه ۲۳.۲، در صورتی که فرایند تصادفی مارکوف باشد رابطه آن با τ خطی است. و برای فرایندهای غیرمارکوف در صورتی که بتوانیم یک طول مارکوف به دست بیاوریم این رابطه خطی به ازای $\tau > l_m$ دیده می‌شود. در شکل زیر ممان‌های شرطی اول و دوم میدان سرعت برای یک سیال بر حسب Δr رسم شده است. باید توجه داشته باشیم که در این مساله خاصیت مارکوف نسبت به r سنجیده شده است. همانطور که می‌بینیم این ممان‌ها در این شکل خط چین عمودی نشان دهنده طول مارکوف l_m است. به ازای $\Delta r > l_m$ رفتار خطی با Δr دارند. [۳۳]



(ب) (آ)

شکل ۵.۲: ممان‌های شرطی اول و دوم در میدان سرعت یک سیال برحسب مقیاس Δr، خط چین عمودی طول مارکوف را نشان می‌دهد. [۳۳]

۵.۲ تخمین طول مارکوف

تا کنون سه روش برای محاسبه طول مارکوف (l_m) معرفی شده است که در این بخش این سه روش یعنی روش χ^2، روش ویلکاکسون[14] و آزمون چپمن-کولموگروف را توضیح خواهیم داد.

۱.۵.۲ روش χ^2

ایده اصلی در این روش مقایسه احتمال توأم سه نقطه‌ای یک فرایند با فرض مارکوف بودن و بدون فرض مارکوف بودن است. به عبارت دیگر احتمال توأم سه نقطه‌ای به شکل کلی زیر را در نظر بگیرید.

$$p\left(x_3, t_3; x_2, t_2; x_1, t_1\right) = p\left(x_3, t_3 | x_2, t_2; x_1, t_1\right) p\left(x_2, t_2; x_1, t_1\right) \qquad (۳۹.۲)$$

حال اگر فرض کنیم فرایند x مارکوف است احتمال توأم سه نقطه‌ای آن را با p_m نشان می‌دهیم و به شکل زیر در به دست می‌آید.

$$p_m\left(x_3, t_3; x_2, t_2; x_1, t_1\right) = p\left(x_3, t_3 | x_2, t_2\right) p\left(x_2, t_2; x_1, t_1\right) \qquad (۴۰.۲)$$

[14] Wilcoxon test

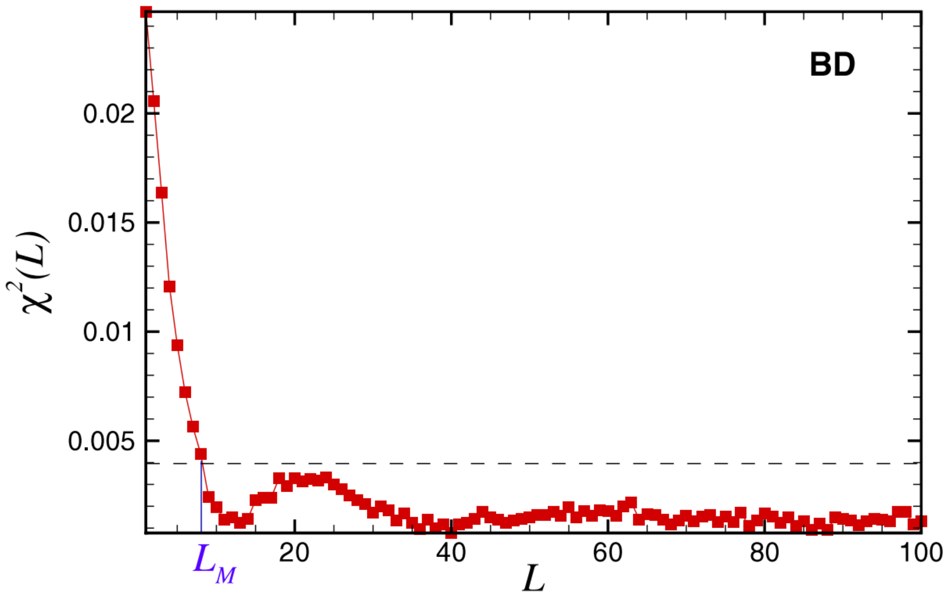
اگر x خاصیت مارکوف را داشته باشد این احتمال‌های توأم باید با هم برابر باشند، برای مقایسه این دو توزیع احتمال توأم از روش χ^2 استفاده می‌کنیم. کمیت χ^2 را در نظر بگیرید که به شکل زیر تعریف می‌شود.

$$\chi^2 = \int dx_3 dx_2 dx_1 \frac{\left[p\left(x_3,t_3;x_2,t_2;x_1,t_1\right) - p_m\left(x_3,t_3;x_2,t_2;x_1,t_1\right)\right]^2}{\left(\sigma_{3j}^2 + \sigma_m^2\right)} \tag{۴۱.۲}$$

که σ_{3j}^2 و σ_m^2 به ترتیب واریانس $p\left(x_3,t_3;x_2,t_2;x_1,t_1\right)$ و $p_m\left(x_3,t_3;x_2,t_2;x_1,t_1\right)$ هستند. برای تخمین t_m از برآورد درست نمایی بیشینه استفاده می‌کنیم که به شکل زیر به دست می‌آید.

$$p\left(t_3 - t_1\right) = \Pi_{x_3,x_2 x_1} \frac{1}{\sqrt{2\pi\left(\sigma_{3j}^2+\sigma_m^2\right)}} \exp\left\{-\frac{\left[p(x_3,t_3;x_2,t_2;x_1,t_1)-p_m(x_3,t_3;x_2,t_2;x_1,t_1)\right]^2}{2\left(\sigma_{3j}^2+\sigma_m^2\right)}\right\} \tag{۴۲.۲}$$

مشخص است برای آن که $p(t_3 - t_1)$ بیشینه باشد، χ_ν^2 باید کمینه بشود که $\chi_\nu^2 = \frac{\chi^2}{N}$ و N تعداد درجات آزادی است. بنابراین جایی که χ_ν^2 کمینه بشود معادل طول ماروکوف $t_m = t_3 - t_1$ است. [۱۵، ۳۴-۳۶]



شکل ۶.۲: معیار χ_ν^2 که طول مارکوف مدل Ballistic deposition را نشان می‌دهد. [۳۵]

۲.۵.۲ روش ویلکاکسون

فرایند تصادفی y و z را در نظر بگیرید که از روی فرایند x ساخته می‌شوند، به این شکل که فرایند y نشان‌دهنده x_3 در زمان t_3 به این شرط که در زمان t_2، x_2 دیده شده باشد. فرایند z نیز نشان‌دهنده x_3 در t_3 به شرطی که x_2 در t_2 و x_1 در t_1 مشاهده شده باشند، به عبارت دیگر داریم.

$$\begin{aligned} y\left(x_2, t_3, t_2\right) &= \left. x_3\left(t_3\right)\right|_{x_2(t_2)} \\ z\left(x_2, x_1, t_1, t_2, t_3\right) &= \left. x_3\left(t_3\right)\right|_{x_2(t_2), x_1(t_1)} \end{aligned} \tag{۴۳.۲}$$

حال اگر $p(y)=\tilde{p}(z)$ باشد می‌گوییم فرایند مارکوف است، اما این دو چگالی احتمال هر دو نامشخص هستند. برای مقایسه این دو چگالی احتمال فرض کنید که نمونه‌های مستقل $y_1, \ldots, y_n$ و $z_1, \ldots, z_m y_n$ را از فرایندهای y و z داریم. حال اگر این داده‌ها را به شکل صعودی مرتب کنیم داریم.

$$y_1 < y_2 < z_1 < y_3 < z_2 < z_3 < y_4 < \cdots \tag{۴۴.۲}$$

برای بررسی فرض $p(y)=\tilde{p}(z)$ از تعداد وارونگی استفاده می‌کنیم، به تعداد y_jهایی که $y_j < z_i$ تعداد وارونگی z_i گفته می‌شود. به عنوان مثال در نمونه بالا تعداد وارونگی‌های z_1 برابر ۲ است و تعداد وارونگی‌های z_2 برابر ۳ است. تعداد کل وارونگی‌ها برای تمام zها را با Q نمایش می‌دهیم. حال اگر $p(y)=\tilde{p}(z)$ باشد، Q توزیع گاوسی دارد که میانگین آن برابر است با

$$\langle Q\rangle_{p=\tilde{p}} = \frac{mn}{2} \tag{۴۵.۲}$$

و واریانس Q به ازای $n, m > 25$ برابر است با

$$\sigma(m,n)^2 = \frac{nm(n+m+1)}{12} \tag{۴۶.۲}$$

انحراف از مقدار چشم‌داشتی ۴۵.۲ معمولاً با قدر مطلق اختلاف $Q - \langle Q \rangle_{p=\tilde{p}}$ بررسی می‌شود.

$$q = |Q - \langle Q \rangle_{p=\tilde{p}}| = q\left(x_2, x_1, t_1, t_2, t_3\right) \tag{۴۷.۲}$$

فرض $p(y) = \tilde{p}(z)$ زمانی پذیرفته است که مقدار q از یک کران مشخص کوچکتر باشد، این کران با استفاده از یک به اصطلاح سطح اهمیت α مشخص می‌شود و معمولاً با q_α نمایش داده می‌شود. برای یک α مشخص، q_α به شکلی تعیین می‌شود که احتمال اینکه q مقداری بزرگ‌تر از q_α داشته باشد برابر α بشود. به عبارت دیگر حتی اگر $p(y) = \tilde{p}(z)$ برقرار باشد هنوز هم به اندازه α این احتمال وجود دارد که q مقداری بزرگ‌تر از q_α داشته باشد. از آنجایی که q قدر مطلق یک متغیر تصادفی با توزیع گاوسی با واریانس به شکل ۴۶.۲ است، کران q_α را می‌توان با استفاده از معکوس تابع خطای گاوسی به دست آورد. [۱۶، ۳۷-۳۹]

۳.۵.۲ آزمون چپمن-کولموگروف

روش سوم که برای تخمین طول مارکوف معرفی خواهد شد به طور مستقیم از معادله چپمن-کولموگروف استفاده خواهد کرد. همان‌طور که قبلاً گفته شد معادله چپمن-کولموگروف ۸.۲ برای فرایندهای مارکوف حتماً برقرار است، اما اگر این معادله برای فرایندی برقرار باشد این فرایند لزوماً مارکوف نیست. بنابراین برقرار بودن معادله چپمن-کولموگروف شرط لازم برای یک فرایند مارکوف است. فرض کنید فرایند x غیرمارکوف است. بنابراین معادله چپمن-کولموگروف به شکلی که برای یک فرایندهای مارکوف برقرار است برای فرایند x برقرار نیست. اما اگر دو نقطه x_i و x_j که نشان دهنده فرایند تصادفی در زمان t_i و t_j هستند را در نظر بگیریم. می‌توانیم برقرار بودن یا نبودن معادله چپمن-کولموگروف را برای این نقاط بررسی کنیم. برای این کار نیاز است که x_k که نشان‌دهنده فرایند تصادفی در t_k که بین t_i و t_j است را در نظر بگیریم. می‌توان با تعریف کمیت S که نشان دهنده اختلاف دو طرف معادله چپمن-کولموگروف است برقرار بودن این معادله را ببینیم.

$$S(x_i, t_i; x_j, t_j) = p(x_i, t_i; x_j, t_j) - \int_{-\infty}^{+\infty} \mathrm{d}x_k p(x_i, t_i|x_k, t_k) p(x_k, t_k; x_j, t_j) \tag{۴۸.۲}$$

در واقع اگر رابطه چپمن-کولموگروف برای x_i و x_j برقرار باشد مقدار S برابر با صفر خواهد بود و اگر این رابطه برقرار نباشد مقدار S مخالف صفر است. برای اینکه بتوانیم معیاری از آزمون چپمن-کولموگروف به ازای همه

مقادیر به دست آوریم می‌توان از رابطه بالا قدرمطلق گرفت و سپس از آن نسبت به x_i و x_j انتگرال گرفت.

$$S(t_i, t_j) = \int dx_i dx_j \left| S(x_i, t_i; x_j, t_j) \right| \tag{۴۹.۲}$$

اگر در زمان‌های مشخص t_i و t_j که $t_i > t_j$، به ازای همه مقادیر x_i و x_j مقدار $S(x_i, t_i; x_j, t_j)$ برابر با صفر باشد، $S(t_i, t_j)$ نیز صفر می‌شود. بنابراین طول مارکوف $l_m = t_i - t_j$ است. ذکر این نکته لازم است که S به ازای همه زمان‌های t_k با این شرط که $t_i > t_k > t_j$ باید صفر باشد. [۱۷]

اگر فرایند مورد نظر یک فرایند مانا باشد یا آن را مانا فرض کنیم می‌توان به جای محاسبه S برای همه مقادیر t_i و t_j آن را بر حسب $\tau = t_i - t_j$ محاسبه کرد. برای جمله دوم رابطه ۴۸.۲ نیز می‌توانیم $\tau' = t_i - t_k$ و $\tau'' = t_k - t_j$ را در نظر بگیریم. بنابراین داریم:

$$S(x_i, x_j; \tau) = p(x_i, x_j; \tau) - \int_{-\infty}^{+\infty} \mathrm{d}x_k p(x_i|x_k; \tau') p(x_k, x_j; \tau'') \tag{۵۰.۲}$$

و معادل رابطه ۴۹.۲ هم برای فرایند مانا می‌توانیم بنویسیم:

$$S\left(\tau\right) = \int dx_i dx_j \left| S(x_i, x_j; \tau) \right| \tag{۵۱.۲}$$

در شرایط عملی که ممکن است فقط یک سری زمانی داشته باشیم استفاده از رابطه ۴۹.۲ برای تخمین طول مارکوف ممکن نیست و باید از رابطه بالا استفاده کنیم. ذکر این نکته لازم است که، در شرایط عملی باید خطای محاسبه S را نیز به دست بیاوریم و در نظر داشته باشیم که طول مارکوف در جایی تعیین می‌شود که S با در نظر گرفتن خطا صفر خواهد شد. [۱۹]

در مسائلی که زمان گسسته است باید دقت کنیم که طول مارکوف برابر با $l_m = t_i - t_j - 1$ خواهد بود به این دلیل که همانطور که دیدیم معادله چپمن-کولموگروف برای یک فرایند مارکوف وقتی برقرار است که $\tau = 2$ است اما طول مارکوف $l_m = 1$ است. نکته مهم دیگر درباره این روش این است که S را به ازای $\tau = t_i - t_j \leq 1$ نمی‌توان محاسبه کرد، بنابراین اگر برای فرایندی S به ازای $\tau = 2$ صفر شد تنها می‌توان نتیجه گرفت که طول مارکوف این فرایند کوچک‌تر از 1 است یعنی این فرایند یا مارکوف است یا مستقل از حالت‌های گذشته است.

محاسبه خطا برای آزمون چپمن-کولموگروف

برای محاسبه خطای S می‌توان از مفهوم انتشار خطا استفاده کرد. در این روش با استفاده از خطای متغیرهای یک تابع می‌توان خطای تابع را محاسبه کرد. بدین منظور اگر $f(x_1, x_2, x_3, \cdots, x_n)$ یک تابع با متغیرهای $x_1, x_2, \cdots, x_n$ باشد، می‌توان با رابطه زیر خطای f را محاسبه کرد.

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \sigma_{x_2}^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \sigma_{x_n}^2} \tag{۵۲.۲}$$

در رابطه بالا σ_f نشان دهنده خطای f و $\sigma x_1, \sigma x_2, \cdots, \sigma x_n$ نشان دهنده خطای متغیرهای $x_1, x_2, \cdots, x_n$ است.

بنابراین برای محاسبه خطای S ابتدا باید خطای توزیع‌های احتمال را به دست آورد. خطای یک توزیع احتمال یا توزیع احتمال شرطی را می‌توان با استفاده از رابطه زیر به دست آورد. [۴۰]

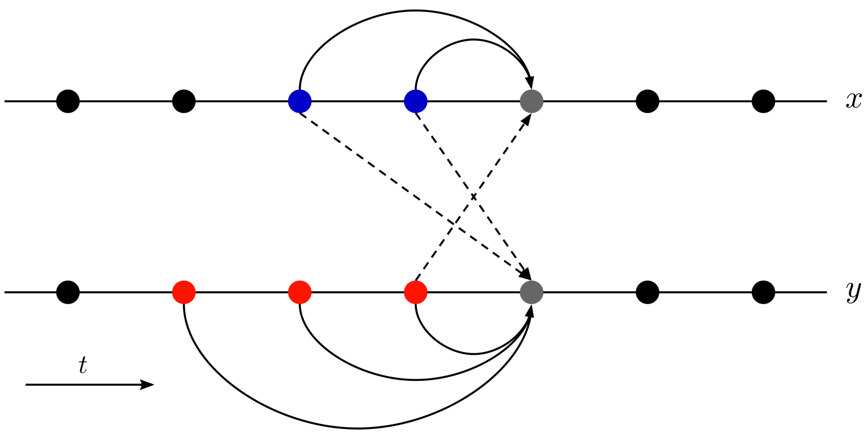
$$\sigma_{p(x)} = \frac{\sqrt{p(x)(1 - p(x)\mathrm{d}x)}}{n\mathrm{d}x} \tag{۵۳.۲}$$

با استفاده از این دو رابطه و رابطه ۴۸.۲ می‌توان خطای S را به شکل زیر به دست آورد.

$$\sigma_{\mathbf{S}} = \sqrt{\sigma_{p_{ij}}^2 + \sigma_{p_{ik}}^2 \times p_{kj}^2 + p_{ik}^2 \times \sigma_{p_{jk}}^2} \tag{۵۴.۲}$$

۶.۲ فرایندهای غیرمارکوف وابسته

دو فرایند تصادفی x و y را در نظر بگیرید که غیر از خودشان به یکدیگر نیز وابسته‌اند که این وابستگی می‌تواند دارای حافظه نیز باشد. شکل زیر نمایشی از وابستگی این دو فرایند به خودشان و به یکدیگر است. به این معنی که گذشته هر یک از فرایندها نه تنها در آینده خودشان تاثیر گذار است بلکه بر آینده فرایند دیگر نیز اثر می‌گذارد.



شکل ۷.۲: دو فرایند وابسته x و y که نسبت به یکدیگر و خودشان حافظه متفاوتی دارند. خطوط کامل حافظه نسبت به خود فرایند و خطوط خط چین حافظه نسبت به یکدیگر را نشان می‌دهند.

می‌توان این دو فرایند را با استفاده از بردار تصادفی بررسی کرد. بردار $\vec{r}$ را به شکل زیر تعریف می‌کنیم.

$$\vec{r} = (x, y) \tag{۵۵.۲}$$

در این شرایط رابطه ۳.۲ را می‌توانیم به شکل زیر بازنویسی کنیم.

$$\begin{aligned} & p(\vec{r}_n, t_n; \vec{r}_{n-1}, t_{n-1} \ldots; \vec{r}_0, t_0) \\ & \quad = p(\vec{r}_n, t_n \mid \vec{r}_{n-1}, t_{n-1} \ldots; \vec{r}_0, t_0) \ldots p(\vec{r}_2, t_2 \mid \vec{r}_1, t_1; \vec{r}_0, t_0) p(\vec{r}_1, t_1 \mid \vec{r}_0, t_0) p(\vec{r}, t_0) \end{aligned} \tag{۵۶.۲}$$

ویژگی مارکوف برای این نوع فرایندها نیز قابل بررسی است. مشابه رابطه ۴.۲ فرایند برداری را مارکوف می‌گویند که رابطه زیر برای آن برقرار باشد.

$$p(\vec{r}_n, t_n \mid \vec{r}_{n-1}, t_{n-1} \ldots; \vec{r}_0, t_0) = p(\vec{r}_n, t_n \mid \vec{r}_{n-1}, t_{n-1}) \tag{۵۷.۲}$$

و در نهایت برای توزیع احتمال توأم می‌توان نوشت.

$$\begin{aligned} &p(\vec{r}_n, t_n; \vec{r}_{n-1}, t_{n-1} \dots ; \vec{r}_0, t_0) \\ &\quad = p(\vec{r}_n, t_n \mid \vec{r}_{n-1}, t_{n-1}) \dots p(\vec{r}_2, t_2 \mid \vec{r}_1, t_1) p(\vec{r}_1, t_1 \mid \vec{r}_0, t_0) p(\vec{r}_0, t_0) \end{aligned} \quad (۵۸.۲)$$

نکته قابل توجه و مهم در اینجا پیدا کردن طول مارکوف برای دو فرایند وابسته است که در این حالت فرایندهای x و y می‌توانند نسبت به خودشان و یکدیگر طول‌های متفاوتی داشته باشند. بدین منظور باید روش پیدا کردن طول مارکوف برای دو فرایند وابسته بازنویسی شوند. که در فصل آینده آزمون چپمن-کولموگروف را برای این فرایندها تعمیم خواهیم داد.

۷.۲ مدل خودبرگشت

برای آزمایش روش چپمن-کولموگروف برای تخمین طول مارکوف دو فرایند وابسته، در ابتدا بهتر است که در قدم اول با داده‌هایی کار کنیم که طول مارکوف آن‌ها را می‌دانیم. به این ترتیب می‌توان برخی از نقاط قوت و ضعف احتمالی را بهتر شناسایی کرد. برای این کار از مدلی به نام مدل خودبرگشت استفاده می‌کنیم. یک مدل خودبرگشت از مرتبه p که آن را با $AR(p)$ نشان می‌دهند به شکل زیر نوشته می‌شود.

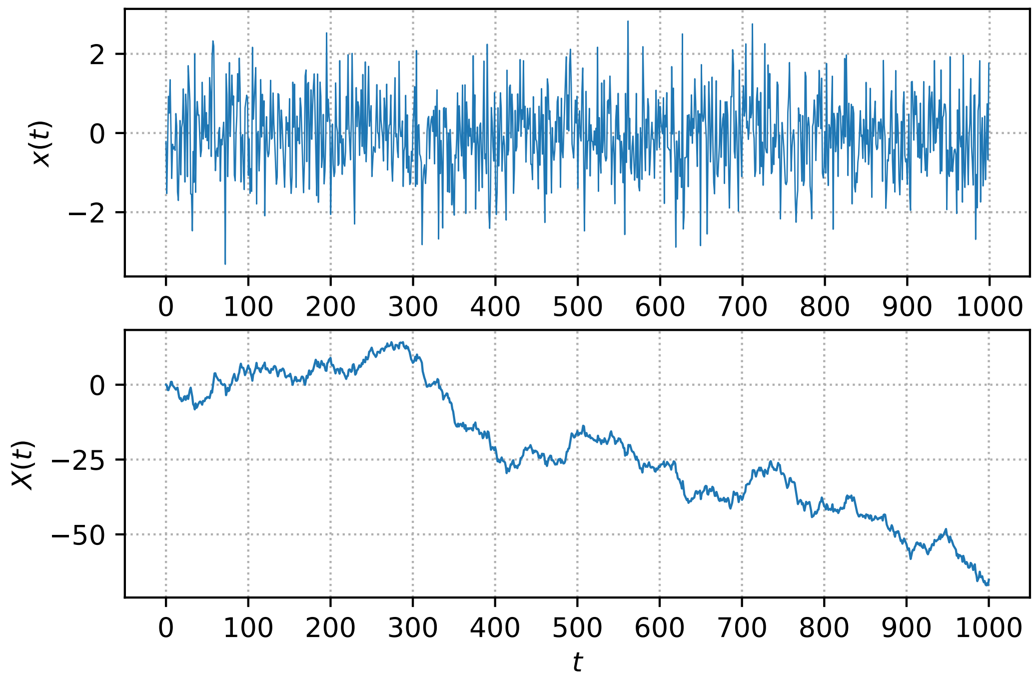
$$x_t = \sum_{n=1}^{p} \phi_n x_{t-n} + \xi_t \quad (۵۹.۲)$$

ضرایب ϕ_n ضرایب ثابت غیر صفر هستند و ξ_t یک نویز گاوسی با میانگین صفر و واریانس σ_ξ^2 است. میانگین x_t در رابطه ۵۹.۲ برابر با صفر است اگر نیاز به ساخت سری زمانی با میانگین غیر صفر باشد کافی است x_t با $x_t - \mu$ جایگزین بشود. [۹]

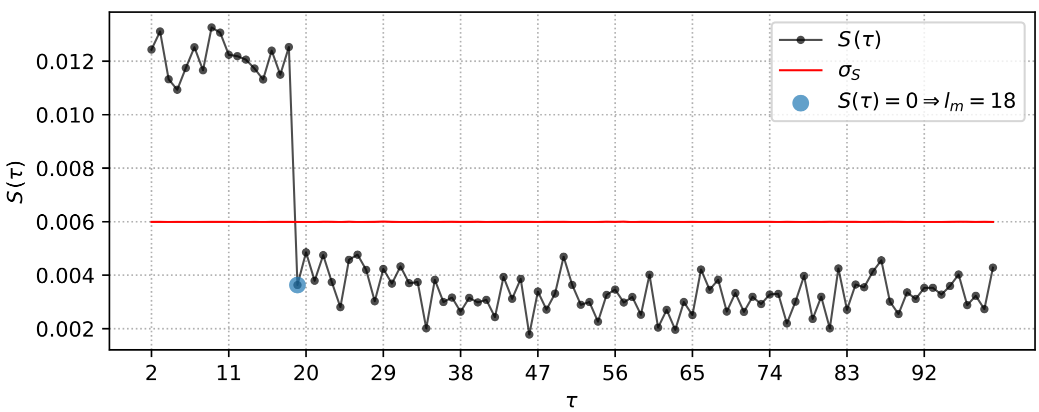
با رابطه بالا می‌توان سری‌های زمانی تولید کرد که به تعداد p قدم به گذشته خودشان وابسته هستند. در شکل ۸.۲ یک نمونه از فرایند تصادفی رسم شده که با رابطه ۵۹.۲ تولید شده است. برای تولید این سری زمانی مقادیر p و ϕ_n به شکل زیر تعیین شده‌اند.

$$p = 18, \phi_n = \frac{0.3}{p}$$

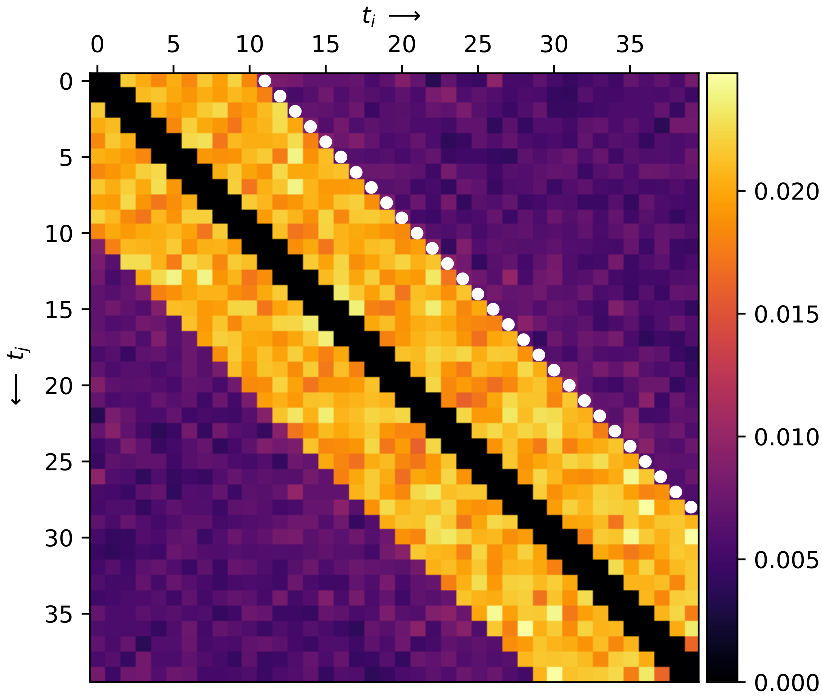
در صورتی که فقط یک نمونه با استفاده از رابطه ۵۹.۲ تولید کنیم، تنها قادر خواهیم بود که آزمون چپمن-کولموگروف را با فرض همگن بودن سری زمانی استفاده کنیم. در شکل ۹.۲ نتیجه تست چپمن-کولموگروف را برای یک نمونه از فرایند ۵۹.۲ را می‌بینیم.



شکل ۸.۲: یک نمونه از فرایند تصادفی که با استفاده از رابطه ۵۹.۲ با $p = 18$ تولید شده است. در این شکل $X_{t+1} = X_t + x_t$ است.



شکل ۹.۲: آزمون چپمن-کولموگروف برای سری زمانی تولید شده با استفاده از ۵۹.۲. مقدار $S(\tau)$ به ازای $\tau = 19$ با در نظر گرفتن خطا صفر شده است که نشان دهنده وابستگی فرایند x به ۱۸ قدم قبلی خودش است.



شکل ۱۰.۲: ماتریس $S(t_i, t_j)$ برای 3.5×10^5 نمونه با $p = 10$.

همانطور که در شکل ۹.۲ مشخص است مقدار $S(\tau)$ به ازای $\tau = 19$ با در نظر گرفتن خطا صفر شده‌است که نشان دهنده این مسأله است گه فرایند x مستقل از حالت خود در ۱۹ قدم قبل است، اما به ۱۸ قدم قبلی خود وابسته است.

برای محاسبه ماتریس $S(t_i, t_j)$ به بیش از یک نمونه برای محاسبه توزیع‌های احتمال شرطی و توأم نیاز داریم. در شکل ۱۰.۲ نتیجه تست چپمن-کولموگروف برروی 3.5×10^5 نمونه رسم شده است. در این نمونه‌ها p برابر با ۱۰ در نظر گرفته شده بود. در شکل ۱۰.۲ نقاط سفید نشان دهنده اوین جایی است که $S(t_i, t_j)$ با در نظر گرفتن خطا صفر شده است، همان طور که پیدا است به ازای $t_i - t_j = 11$ مقدار $S(t_i, t_j)$ با در نظر گرفتن خطا صفر شده است. که نشان دهنده تخمین درست $p = 10$ است.

فصل ۳

تعمیم آزمون چپمن-کولموگروف

۱.۳ آزمون چپمن-کولموگروف برای دو فرایند وابسته

برای تخمین طول مارکوف با استفاده از آزمون چپمن-کولموگروف ابتدا باید تعمیم معادله چپمن-کولموگروف ۸.۲ برای دو فرایند وابسته را در نظر داشته باشیم.

$$p(\vec{r}_{i+2},\vec{r}_i)=\int_{-\infty}^{+\infty}\mathrm{d}\vec{r}_{i+1}p(\vec{r}_{i+2}|\vec{r}_{i+1})p(\vec{r}_{i+1},\vec{r}_i) \tag{۱.۳}$$

مشابه رابطه ۴۸.۲ کمیت S را نیز می‌توان به شکل زیر به دست آورد.

$$S(\vec{r}_i,t_i;\vec{r}_j,t_j)=p(\vec{r}_i,t_i;\vec{r}_j,t_j)-\int_{-\infty}^{+\infty}\mathrm{d}\vec{r}_k p(\vec{r}_i,t_i|\vec{r}_k,t_k)p(\vec{r}_k,t_k;\vec{r}_j,t_j) \tag{۲.۳}$$

در حالتی که چند فرایند وابسته داریم، هر یک از فرایندها می‌توانند نسبت به خودشان و یکدیگر طول‌های مارکوف متفاوتی داشته باشند. بنابراین وقتی دو فرایند وابسته داریم ممکن است تا ۴ طول مارکوف متفاوت داشته باشیم، برای

پیدا کردن این طول‌ها می‌توانیم روابط زیر را در نظر بگیریم.

$$\begin{aligned}
S_1(x_i,t_i;x_j,t_j) &= \textstyle\int_{-\infty}^{+\infty} \mathrm{d}y_i \mathrm{d}y_j S_{ij}(\vec{r}_i,t_i;\vec{r}_j,t_j)\\
S_2(x_i,t_i;y_j,t_j) &= \textstyle\int_{-\infty}^{+\infty} \mathrm{d}y_i \mathrm{d}x_j S_{ij}(\vec{r}_i,t_i;\vec{r}_j,t_j)\\
S_3(y_i,t_i;x_j,t_j) &= \textstyle\int_{-\infty}^{+\infty} \mathrm{d}x_i \mathrm{d}y_j S_{ij}(\vec{r}_i,t_i;\vec{r}_j,t_j)\\
S_4(y_i,t_i;y_j,t_j) &= \textstyle\int_{-\infty}^{+\infty} \mathrm{d}x_i \mathrm{d}x_j S_{ij}(\vec{r}_i,t_i;\vec{r}_j,t_j)
\end{aligned} \tag{۳.۳}$$

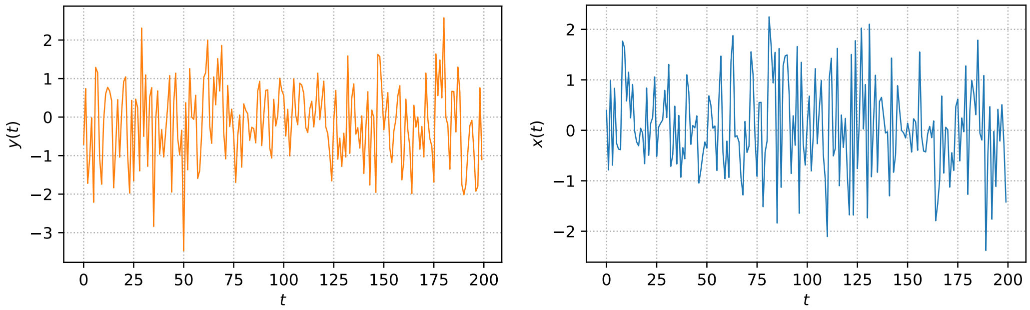
روابط بالا به ترتیب از بالا برای تخمین طول مارکوف فرایند x نسبت به خودش، فرایند x نسبت به y، فرایند y نسبت به x و فرایند y نسبت به خودش استفاده می‌شوند. در نهایت از قدرمطلق روابط بالا انتگرال خواهیم گرفت.

$$\begin{aligned}
S_1(t_i,t_j) &= \int dx_i dx_j \left|S_1(x_i,t_i;x_j,t_j)\right|, \qquad S_1(t_i,t_j)=0 \Longrightarrow l_1 = t_i - t_j\\
S_2(t_i,t_j) &= \int dx_i dy_j \left|S_2(x_i,t_i;y_j,t_j)\right|, \qquad S_2(t_i,t_j)=0 \Longrightarrow l_2 = t_i - t_j\\
S_3(t_i,t_j) &= \int dy_i dx_j \left|S_3(y_i,t_i;x_j,t_j)\right|, \qquad S_3(t_i,t_j)=0 \Longrightarrow l_3 = t_i - t_j\\
S_4(t_i,t_j) &= \int dy_i dy_j \left|S_4(y_i,t_i;y_j,t_j)\right|, \qquad S_4(t_i,t_j)=0 \Longrightarrow l_4 = t_i - t_j
\end{aligned} \tag{۴.۳}$$

طول مارکوف فرایند x نسبت به خودش با l_1 نشان داده می‌شود. طول‌های مارکوف دیگر نیز با l_2، l_3 و l_4 نشان داده می‌شوند. در این حالت نیز می‌توانیم برای دو فرایند غیرمارکوف وابسته مانا آزمون چپمن-کولموگروف را بر حسب τ بنویسیم.

تا اینجا آزمون چپمن-کولموگروف را برای دو فرایند غیرمارکوف وابسته تعمیم داده‌ایم، حال برای آزمایش این روش نیاز به مدلی داریم که بتواند سری‌های زمانی وابسته با طول مارکوف دلخواه تولید کند. در بخش بعدی با تعمیم مدل خودبرگشت مدل ساده‌ای را معرفی می‌کنیم که می‌توان طول مارکوف دو فرایند x و y را نسبت به خودشان و نسبت به یکدیگر در آن مشخص کرد.

۲.۳ دو فرایند غیرمارکوف وابسته با طول مارکوف دلخواه



(آ) فرایند x با $l_1 = 20$، $l_2 = 10$.

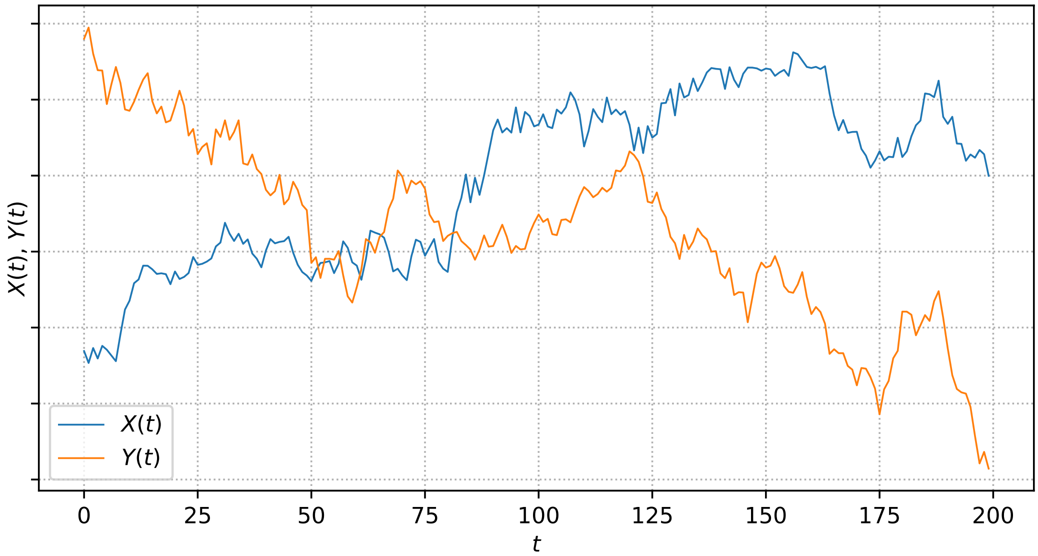
(ب) فرایند y با $l_3 = 15$، $l_4 = 25$.

شکل ۱.۳: یک نمونه از فرایند x و فرایند y تولید شده با استفاده از ۵.۳.

در مدل خودبرگشت ۵۹.۲ دیدیم که برای ساخت فرایند تصادفی با حافظه دلخواه حالت‌های گذشته x در حالت بعدی x به شکل مجموع آن‌ها دخالت داده شده‌اند. برای ایجاد جفت‌شدگی بین دو فرایند تصادفی x و y نیز می‌توان حالت‌های گذشته هر یک از فرایندها را در دیگری دخالت دهیم. برای این کار مانند بخش قبل حالت‌های گذشته y را به شکل وزن دار با رابطه ۵۹.۲ جمع می‌کنیم. فرایند دیگری هم مشابه x با ضرایب مخصوص به خود برای y می‌نویسیم. بنابراین داریم:

$$
\begin{aligned}
x_{t+1} &= \sum_{n=1}^{l_1} \phi_n^{xx} x_{t-n} + \sum_{n=1}^{l_2} \phi_n^{xy} y_{t-n} + \xi_t^x \\
y_{t+1} &= \sum_{n=1}^{l_3} \phi_n^{yx} x_{t-n} + \sum_{n=1}^{l_4} \phi_n^{yy} y_{t-n} + \xi_t^y
\end{aligned}
\tag{۵.۳}
$$

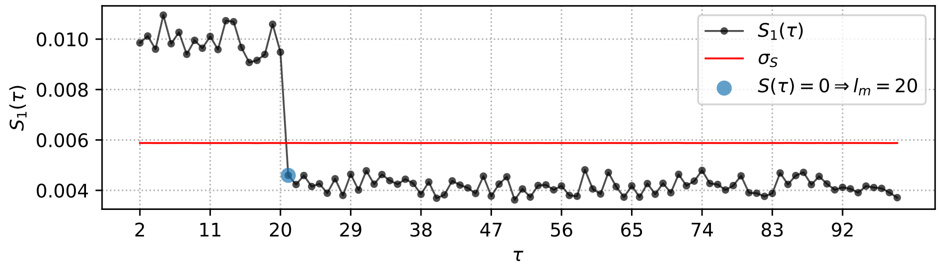
در روابط ۵.۳ چهار ضریب ϕ_n^{xx}، ϕ_n^{xy}، ϕ_n^{yx} و ϕ_n^{yy} ضرایب ثابتی هستند. ξ_t^x و ξ_t^y هم نیروی‌های تصادفی گاوسی مستقل از هم هستند. l_1 و l_2 به ترتیب طول حافظه x نسبت به خودش و y را تعیین می‌کنند، به طور مشابه و l_3 و l_4 نیز طول حافظه y نسبت به فرایند x و خودش را مشخص می‌کنند. در شکل ۱.۳ بخشی از دو فرایند وابسته با $l_1 = 20$، $l_2 = 10$، $l_3 = 15$ و $l_4 = 25$ رسم شده است. در شکل ۲.۳ نیز دو فرایند $X(t)$ و $Y(t)$ رسم شده‌اند که به رابطه ۶.۳ تعریف می‌شوند.



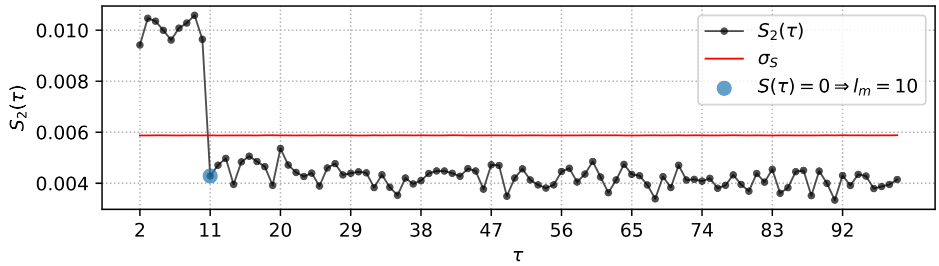
شکل ۲.۳: در این شکل $X(t)$ و $Y(t)$ از رابطه ۶.۳ محاسبه شده است.

$$
\begin{aligned}
X(t+\tau) &= X(t) + \tau x(t) \\
Y(t+\tau) &= Y(t) + \tau y(t)
\end{aligned}
\tag{۶.۳}
$$

در حالتی که فقط یک نمونه از فرایند x و y را داشته باشیم، مثل حالت قبل فقط می‌توانیم S را بر حسب τ به دست بیاوریم. با این تفاوت که برای دو فرایند وابسته باید $S_1(\tau)$، $S_2(\tau)$، $S_3(\tau)$ و $S_4(\tau)$ را محاسبه کنیم. در شکل ۳.۳آ وابستگی فرایند x به خودش با $S_1(\tau)$ و در شکل ۳.۳ب وابستگی فرایند x به y با $S_2(\tau)$ نشان داده شده است. در شکل ۳.۳آ مقدار $S_1(\tau)$ در $\tau = 21$ با در نظر گرفتن خطا صفر شده است که نشان دهنده این است که $l_1 = 20$ است و این مقدار به درستی تخمین زده شده است. در شکل ۳.۳ب نیز با در نظر گرفتن خطا $S_2(\tau)$ در $\tau = 11$ صفر شده است که نشان دهنده تخمین درست برای $l_2 = 10$ است. به طور مشابه در شکل ۴.۳ب نشان دهنده وابستگی فرایند y به فرایند x است که $S_3(\tau)$ با در نظر گرفتن خطا در $\tau = 16$ صفر شده است که نشان دهنده این است که $l_3 = 15$ است. در شکل ۴.۳آ نیز $S_4(\tau)$ در $\tau = 26$ با در نظر گرفتن خطا صفر شده است که نشان دهنده این است که $l_4 = 25$ است.

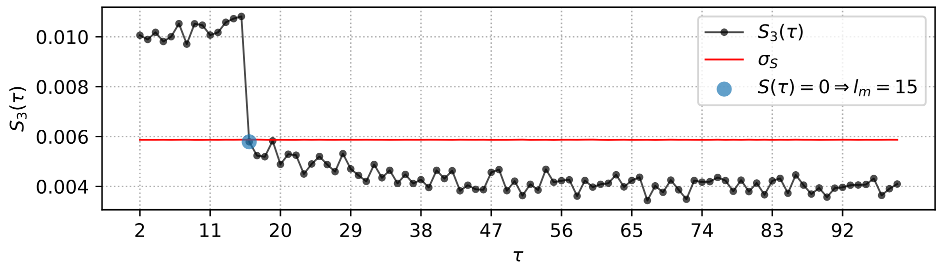


(آ) $S_1(\tau)$ که نشان‌دهنده وابستگی فرایند x به خودش است در $\tau = 21$ با در نظر گرفتن خطا صفر شده است.

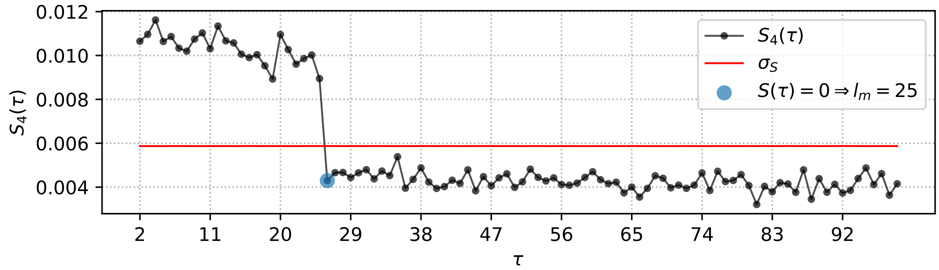


(ب) $S_2(\tau)$ که نشان‌دهنده وابستگی فرایند x به y است در $\tau = 11$ با در نظر گرفتن خطا صفر شده است.

شکل ۳.۳: نمودار $S_1(\tau)$ و $S_2(\tau)$ که نشان دهنده وابستگی فرایند x به خودش و به فرایند y هستند.

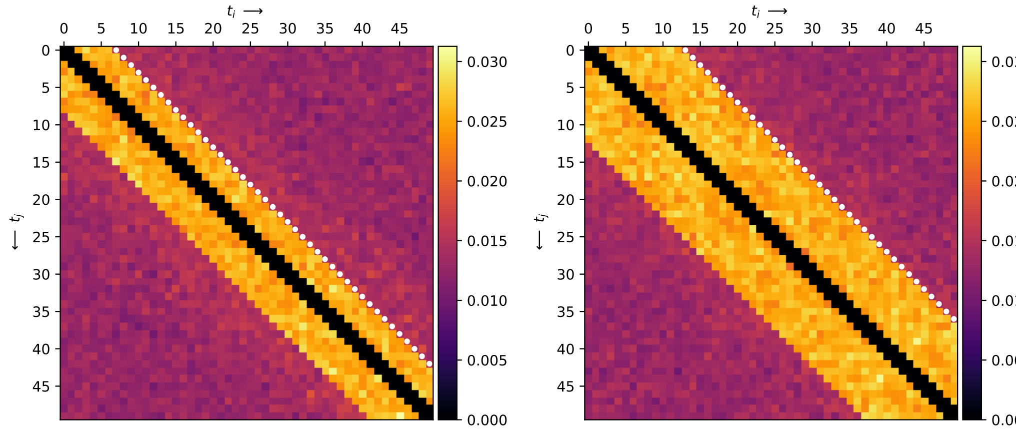


(آ) $S_3(\tau)$ که نشان‌دهنده وابستگی فرایند y به فرایند x است در $\tau = 16$ با در نظر گرفتن خطا صفر شده است.



(ب) $S_4(\tau)$ که نشان‌دهنده وابستگی فرایند y به خودش است در $\tau = 26$ با در نظر گرفتن خطا صفر شده است.

شکل ۴.۳: نمودار $S_3(\tau)$ و $S_4(\tau)$ که نشان دهنده وابستگی فرایند y به فرایند x و به خودش هستند.



(آ) $S_1(t_i, t_i - 13) = 0$، $l_1 = 12$

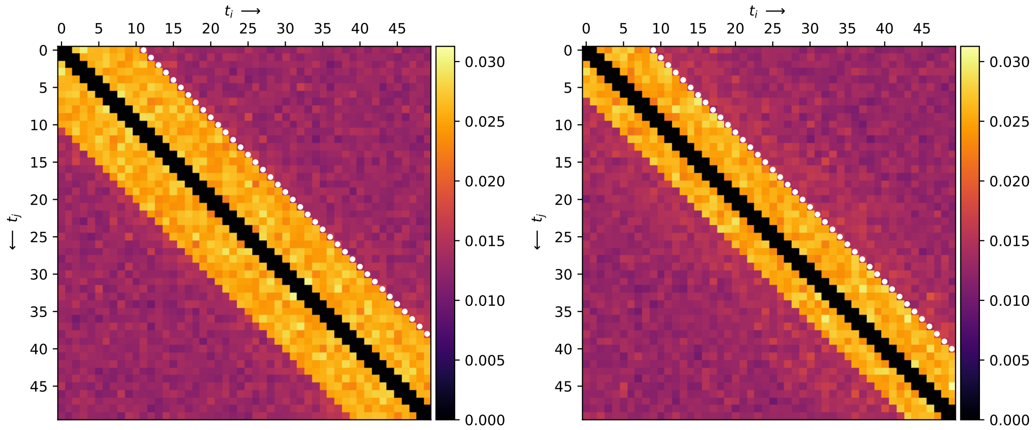
(ب) $S_2(t_i, t_i - 7) = 0$، $l_2 = 6$

شکل ۵.۳: نمودار $S_1(t_i, t_j)$ و $S_2(t_i, t_j)$ که نشان دهنده وابستگی فرایند x به خودش و به فرایند y هستند.

در قدم بعدی آزمون چپمن-کولموگروف را برای بیشتر از یک نمونه از فرایند x و y به کار خواهیم برد. برای این کار 3×10^5 نمونه برای هر یک از فرایندهای x و y تولید شده است. طول مارکوف برای این دو فرایند به شکل زیر تعیین شده است.

$$\begin{aligned} & l_1 = 12, l_2 = 6 \\ & l_3 = 8, l_4 = 10 \end{aligned} \tag{۷.۳}$$

مانند حالتی که یک فرایند داشتیم در اینجا نیز می‌توانیم ماتریس $S(t_i, t_j)$ را به دست بیاوریم که در واقع شامل چهار ماتریس S_1، S_2، S_3 و S_4 می‌شود. شکل ۵.۳ و ۶.۳ نتیاج آزمون چپمن-کولموگروف برای دو فرایند x و y در این تصاویر هم نقاط سفید رنگ نشان دهنده جایی است که S با در نظر گرفتن خطا صفر شده است. شکل ۵.۳آ نمایش گرافیکی ماتریس S_1 است که همانطور که از شکل پیداست $S_1(t_i, t_i - 13) = 0$ که نشان دهنده تخمین درست $l_1 = 12$ است. شکل ۵.۳ب نیز نمایش گرافیکی ماتریس S_2 است که $S_2(t_i, t_j)$ به ازای $t_i \geq t_j + 7$ برابر با صفر است که نشان دهنده این است که $l_2 = 6$ است. به طور مشابه در شکل ۶.۳آ و ۶.۳ب ماتریس‌های S_3 و S_4 به شکل گرافیکی ترسیم شده‌اند. همانطور که پیداست $S_3(t_i, t_j)$ به ازای $t_i \geq t_j + 9$ و $S_4(t_i, t_j)$ به ازای $t_i \geq tj + 11$ با در نظر گرفتن خطا صفر شده‌اند که نشان‌دهنده این است که $l_3 = 8$ و $l_4 = 10$ است که همان مقادیر اولیه ۷.۳ هستند.



(آ) $S_3(t_i, t_i - 9) = 0$، $l_3 = 8$

(ب) $S_4(t_i, t_i - 11) = 0$، $l_4 = 10$

شکل ۶.۳: نمودار $S_3(t_i, t_j)$ و $S_4(t_i, t_j)$ که نشان دهنده وابستگی فرایند y به فرایند x و به خودش هستند.

تا اینجا آزمون چپمن-کولموگروف را برای دو فرایند غیر مارکوف وابسته تعمیم داده‌ایم و همچنین با تعمیم مدل خودبرگشت به دو فرایند، آزمون چپمن-کولموگروف تعمیم یافته را روی سری‌های زمانی تولید شده با این مدل آزمایش کردیم. در فصل بعد از آزمون چپمن-کولموگروف تعمیم یافته برای محاسبه طول مارکوف سری‌های زمانی مالی[1] استفاده خواهیم کرد.

[1] Financial time series

فصل ۴

آزمون چپمن-کولموگروف برای بازارهای مالی

از میان سیستم‌هایی که در زندگی انسان تاثیر گذار هستند، بازارهای مالی یکی از پیچیده‌ترین رفتارها را دارند و از آن‌ها به عنوان یک سیستم پیچیده نامبرده می‌شود. یکی از مباحث مهم در بازارهای مالی پیش‌بینی آینده یک سهم یا شاخص در بازارهای مالی است. رفتار تصادفی بازارهای مالی باعث شده است که بعضی از تحلیل‌گران بازارهای مالی از فرایندهای تصادفی برای تحلیل بازار استفاده کنند. [۱] رفتار بازارهای مالی در واقع برآمده از رفتار جمعی معامله کنندگان بازارهای مالی است که رفتار معامله کنندگان خود وابسته به عوامل بلند مدت و کوتاه مدت است. عوامل بلند مدتی همچون سیاست‌های کلان اقتصادی یک کشور می‌تواند بر رفتار یک بازار مالی در بلند مدت تاثیر بسزایی داشته باشد. بنابراین دور از ذهن نیست که رفتار بازارهای مالی با زمان تغییراتی داشته باشد بنابراین برا محاسبه طول مارکوف بهتر است که این موضوع در نظر گرفته شود.

۱.۴ آماده سازی داده‌ها

برای انجام آزمون چپمن-کولموگروف روی داده‌های بازارهای مالی دو رمزارز [۱] از بازار رمزها انتخاب شده است. به طور دقیق‌تر برای محاسبه طول مارکوف از قیمت قراردادهای دائمی [۲] صرافی بیتمِکس [۳] برای رمز ارز بیتکوین [۴] و

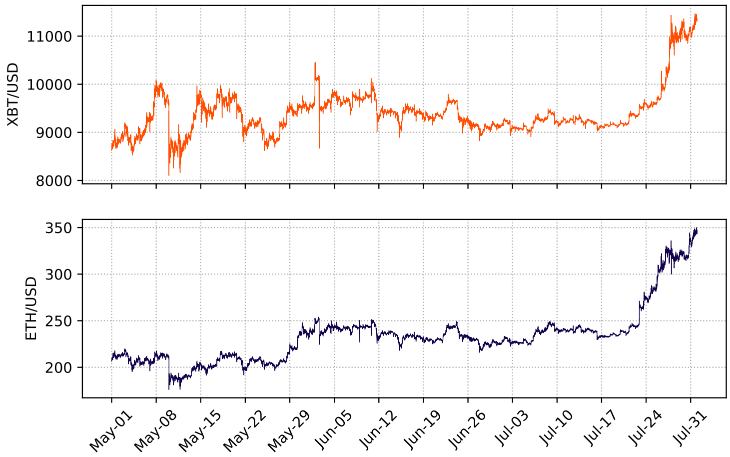
[1] Cryptocurrency

[2] Perpetual contracts

[3] Bitmex

[4] Bitcoin

اتریوم [5] استفاده شده است. [۴۱] داده‌های مورد استفاده داده‌های تک‌تک معاملات قراردادهای دائمی این دو رمز ارز است که با نماد XBT/USD و ETH/USD شناخته می‌شوند. این داده‌ها متعلق به بازه سه ماهه از ابتدای ماه می ۲۰۲۰ تا انتهای جولای ۲۰۲۰ است. با استفاده از قیمت تک‌تک معاملات در نهایت قیمت پایانی برای این قراردادها در هر ثانیه محاسبه شد که در شکل ۱.۴ روند کلی این دو رمزارز در بازه سه ماهه ماه می تا جولای نشان داده شده است.



شکل ۱.۴: قیمت قراردادهای دائمی بیتکوین و اتریوم بر حسب دلار آمریکا در صرافی بیتمکس در بازه سه ماهه می تا جولای ۲۰۲۰.

برای محاسبه طول مارکوف با استفاده از آزمون چپمن-کولموگروف از بازده لگاریتمی [6] استفاده شده است. فرض کنید $x(t)$ یک شاخص یا قیمت یک سهم در یک بازار مالی باشد بازده لگاریتمی به شکل زیر محاسبه می‌شود.

$$r_x(t) = \ln\left(\frac{x(t+1)}{x(t)}\right) \tag{۱.۴}$$

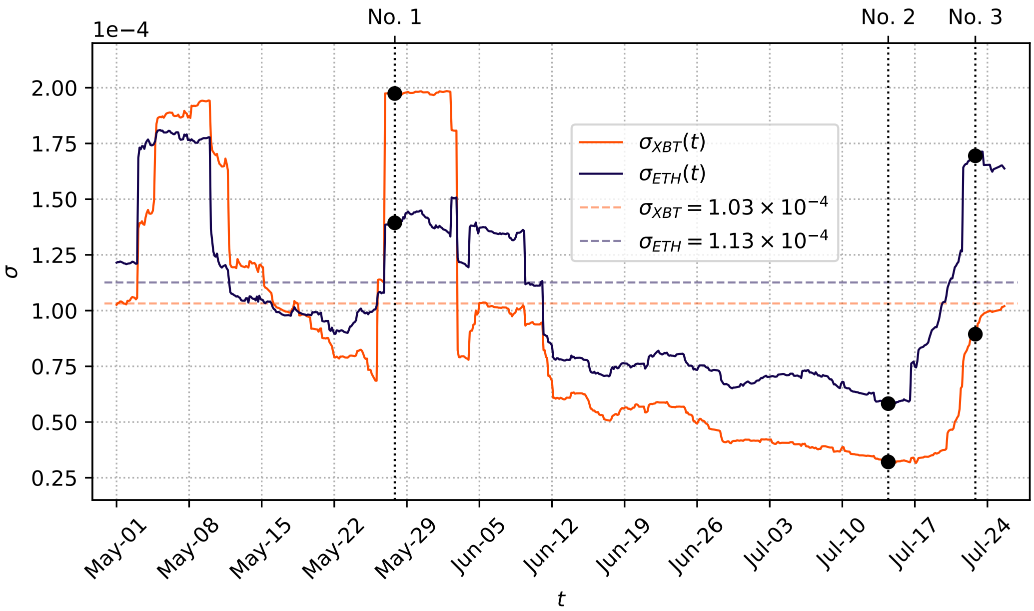
در رابطه بالا $r_x(t)$ نشان دنده بازده لگاریتمی است. دلیل استفاده از بازده لگاریتمی این است که در قدم اول $x(t+1)/x(t)$ نسبت تغییرات قیمت را نسبت به زمان قبل می‌دهد و با گرفتن لگاریتم از این مقادیر می‌توانیم نسبت تغییرات را حول صفر داشته باشیم. در واقع با محاسبه بازده لگاریتمی می‌توانیم داده‌هایی با میانگین نزدیک به صفر داشته باشیم که می‌توانیم آن‌ها را مانا فرض کنیم.

[5] Ethereum

[6] Logarithmic return

۲.۴ محاسبه طول مارکوف داده‌های مالی

همانطور که قبلا اشاره شده ممکن است رفتار بازارهای مالی با زمان تغییر کند و محاسبه طول مارکوف برای بازه‌های زمانی طولانی مثلا سالانه یا ماهانه چندان مفید به نظر نمی‌رسد. بنابراین بهتر است که تخمین طول مارکوف را در بازه‌های کوتاه‌تر مثل هفتگی یا روزانه انجام داد اما از طرفی باید در نظر داشت که هرچه بازه زمانی کوتاه‌تر باشد تعداد نقاط سری‌های زمانی هم کمتر خواهد شد و توزیع چگالی‌های به دست آمده از این سری‌های زمانی خطای زیادی خواهد داشت. در این تحقیق از بازه زمانی هفت روزه استفاده شده است که در حدود 6×10^5 ثانیه است. در ادامه نتیجه آزمون چپمن-کولموگروف برای چند بازه هفت روزه آورده شده است. شکل زیر هم نمودار انحراف معیار بازده‌های لگاریتمی [۷] $r_x(t)$ و $r_y(t)$ است که $r_x(t)$ بازده لگاریتمی بیتکوین و $r_y(t)$ بازده لگاریتمی اتریوم است. در شکل زیر $\sigma_{XBT}(t)$ نشان دهنده انحراف معیار بیتکوین در بازه هفتگی است که ابتدای آن در t قرار دارد و σ_{XBT} نشان دهنده انحراف معیار کل بازه سه ماهه است. کمیت‌های مشابه برای اتریوم با $\sigma_{ETH}(t)$ و σ_{ETH} نشان داده شده‌اند. بازه‌های هفت روزه در طول بازه سه ماهه از می تا جولای ۲۰۲۰ است که به عنوان معیاری از بزرگی نوسان قیمت در نظر گرفته شده است.



شکل ۲.۴: نمودار انحراف معیار بازده‌های لگاریتمی $r_x(t)$ و $r_y(t)$ در بازه‌های هفت روزه در طول بازه سه ماهه از می تا جولای ۲۰۲۰.

[7] Volatility

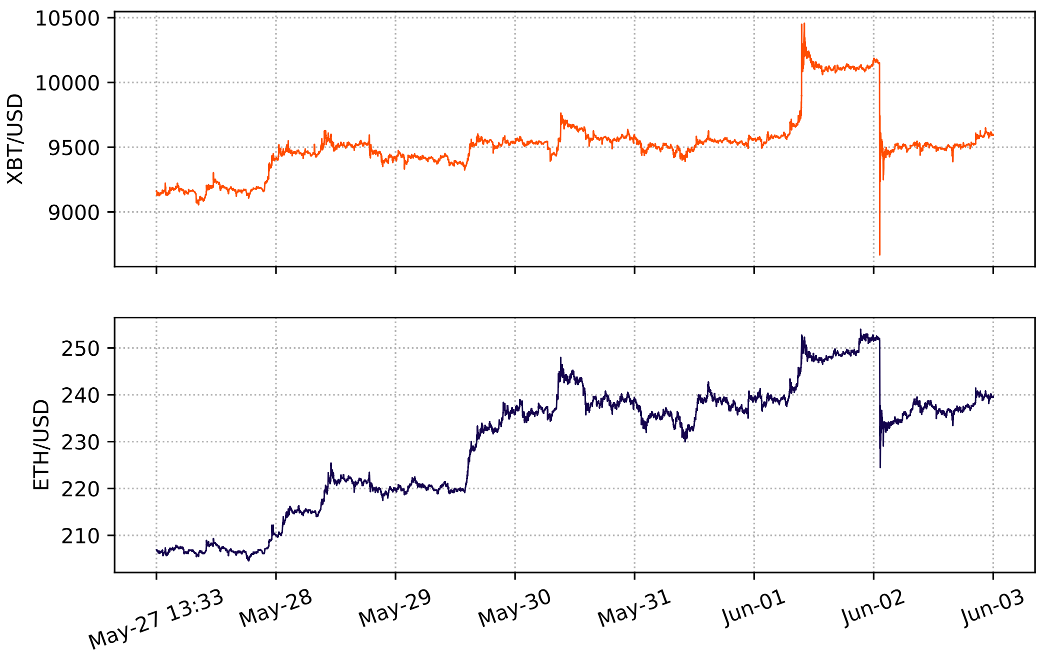
به عنوان نمونه، برای سه بازه هفتگی زیر نتایج تخمین طول ماروکوف آورده شده است که در ادامه به آن می‌پردازیم. مقدار انحراف معیار در این بازه‌ها در شکل ۲.۴ با شماره‌های ۱ تا ۳ نشان داده شده‌اند.

جدول ۱.۴: سه بازه هفتگی که برای محاسبه طول ماروکوف استفاده شده‌اند.

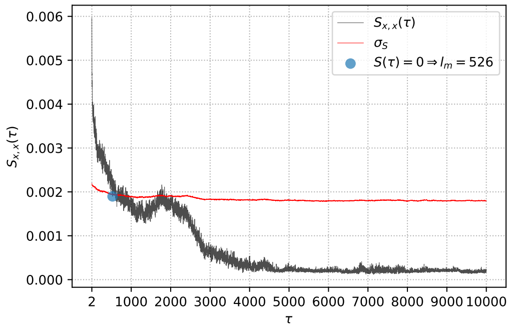
	بازه	$\sigma_{XBT}(t)$	$\sigma_{ETH}(t)$
۱	۲۷ می ۱۳:۳۳:۲۰ تا ۳ جون ۱۳:۳۳:۲۰	1.97×10^{-4}	1.39×10^{-4}
۲	۱۳ جولای ۱۹:۳۴:۱۳ تا ۲۰ جولای ۱۹:۳۴:۱۳	3.22×10^{-5}	5.67×10^{-5}
۳	۲۲ جولای ۲:۵۳:۲۰ تا ۲۹ جولای ۲:۵۳:۲۰	8.93×10^{-5}	1.69×10^{-4}

محاسبه طول مارکوف در بازه شماره ۱

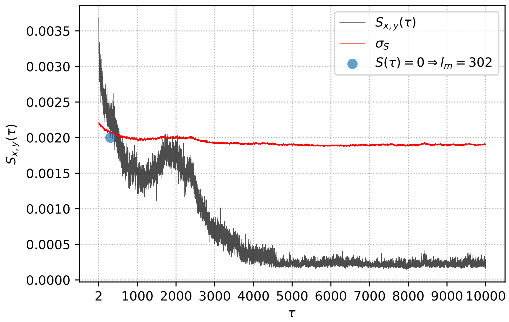
همانطور که در جدول ۱.۴ آمده است بازه شماره ۱ از تاریخ ۲۷ می ۲۰۲۰ ساعت ۱۳:۳۳:۲۰ تا ۳ جون ۲۰۲۰ و تا همان ساعت است. مقدار انحراف معیار بیتکوین در این بازه برابر 1.97×10^{-4} و انحراف معیار اتریوم برابر با 1.39×10^{-4} است. نوسان قیمت بیتکوین و اتریوم در این بازه در شکل زیر آورده شده است. دلیل انتخاب این بازه این است که مقدار انحراف معیار بیتکوین یکی بیشترین مقادیر در بین بازه‌های دیگر را دارد.



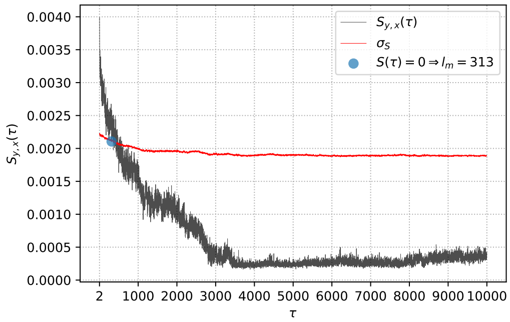
شکل ۳.۴: قیمت قراردادهای دائمی بیتکوین و اتریوم بر حسب دلار آمریکا در بازه شماره ۱.



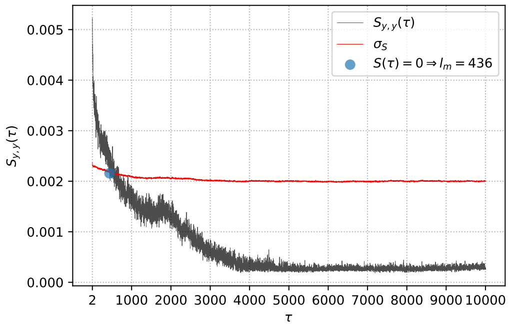
(آ) در این شکل S_1 نشان دهنده طول مارکوف بیتکوین به خودش است، $l_1 = 526$.



(ب) در این شکل S_2 نشان دهنده طول مارکوف بیتکوین به اتریوم است، $l_2 = 302$.



(ج) در این شکل S_3 نشان دهنده طول مارکوف اتریوم به بیتکوین است، $l_3 = 313$.



(د) در این شکل S_4 نشان دهنده طول مارکوف اتریوم به خودش است، $l_4 = 436$.

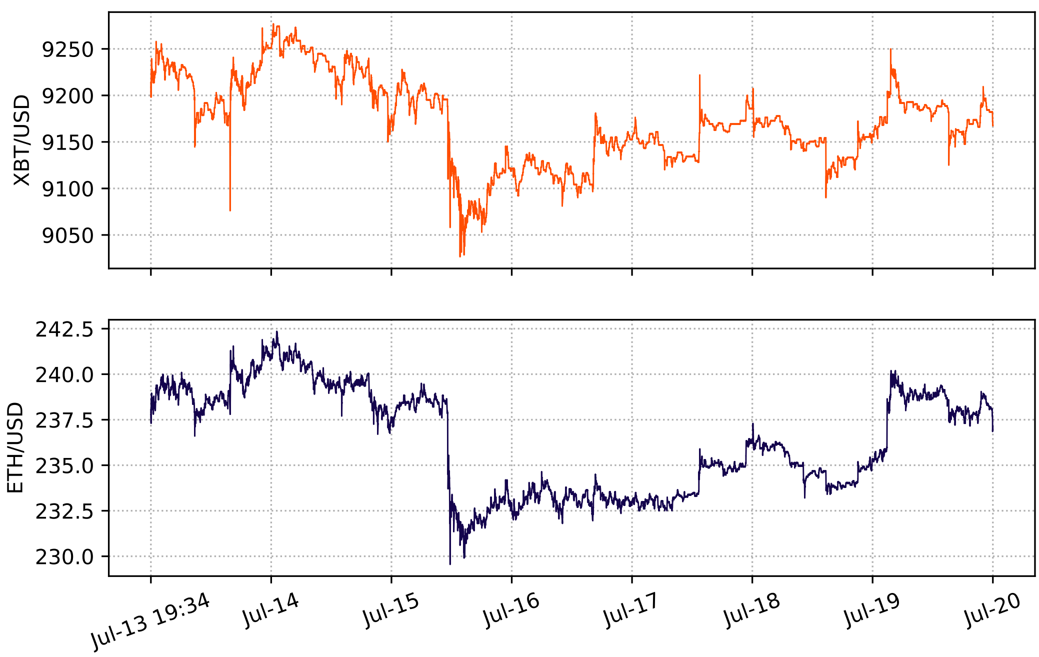
شکل ۴.۴: نمودار آزمون چپمن-کولموگروف برای بیتکوین و اتریوم در بازه شماره ۱.

در شکل ۴.۴ مقادیر $S_1(\tau)$، $S_2(\tau)$، $S_3(\tau)$ و $S_4(\tau)$ برحسب τ کشیده شده‌اند این مقادیر به ترتیب نشان دهنده وابستگی نوسان قیمت بیتکوین به خودش و اتریوم و نوسان قیمت اتریوم به بیتکوین و خودش هستند. همانطور که در شکل ۴.۴آ پیداست مقدار $S_1(\tau)$ در $\tau = 527$ با در نظر گرفتن خطا صفر شده است. از شکل ۴.۴ب نیز پیداست که $S_2(\tau)$ در $\tau = 303$ با در نظر گرفتن خطا صفر شده است. بنابراین برای این بازه $l_1 = 526$ ثانیه و $l_2 = 302$ ثانیه به دست می‌آیند که طول مارکوف بیتکوین نسبت به خودش و اتریوم در این بازه هستند. در شکل ۴.۴د طول مارکوف اتریوم نسبت به خودش یعنی l_4 برابر با 436 به دست آمده است و در شکل ۴.۴ج l_3 که طول مارکوف اتریوم نسبت به بیتکوین است برابر با 313 به دست آمده است.

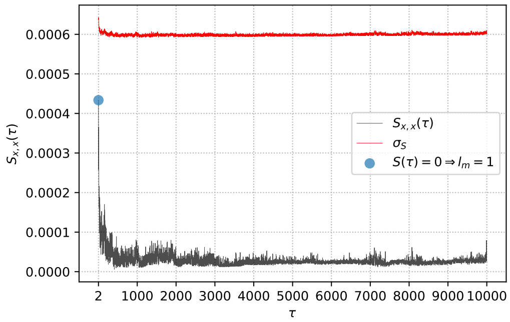
محاسبه طول مارکوف در بازه شماره ۲

بازه شماره ۲ از تاریخ ۱۳ جولای ۲۰۲۰ ساعت ۱۹:۳۴:۱۳ تا ۲۰ جولای ۲۰۲۰ و تا همان ساعت است. مقدار انحراف معیار بیتکوین در این بازه برابر 3.22×10^{-5} و انحراف معیار اتریوم برابر با 5.67×10^{-5} است. این بازه به این دلیل انتخاب شده است که انحراف معیار بیتکوین در این بازه کمینه شده است.

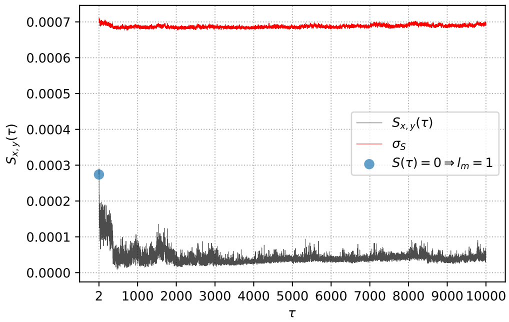
در شکل ۶.۴ نتیجه آزمون چپمن-کولموگروف در این بازه نشان داده شده است. با توجه به شکل‌های ۶.۴آ، ۶.۴ب و ۶.۴ج مقادیر $S_1(\tau)$، $S_2(\tau)$ و $S_3(\tau)$ در $\tau = 2$ با در نظر گرفتن خطا صفر شده‌اند که این مقدار τ با توجه به ۴.۳ نشان دهنده این است که طول مارکوف بیتکوین نسبت به خودش و اتریوم و همچنین اتریوم نسبت به بیتکوین در این بازه کمتر از 1 است. با توجه به شکل ۶.۴د تنها طول مارکوف اتریوم نسبت به خودش برابر با 5 ثانیه است.



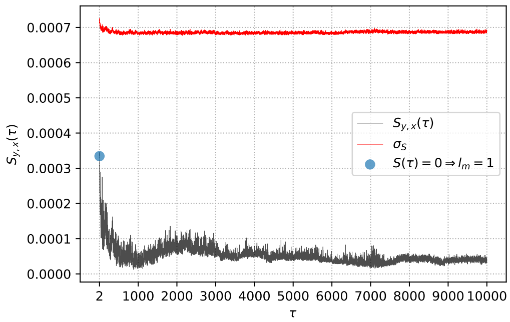
شکل ۵.۴: قیمت قراردادهای دائمی بیتکوین و اتریوم بر حسب دلار آمریکا در بازه شماره ۲.



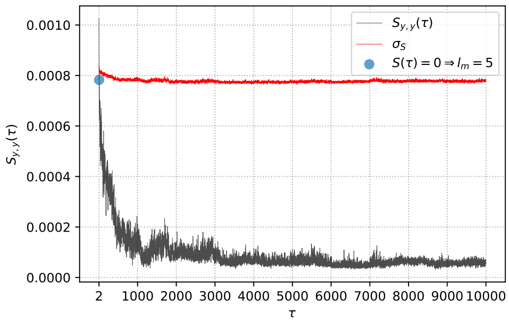
(آ) در این شکل S_1 نشان دهنده طول مارکوف بیتکوین به خودش است، $l_1 \leq 1$.



(ب) در این شکل S_2 نشان دهنده طول مارکوف بیتکوین به اتریوم است، $l_2 \leq 1$.



(ج) در این شکل S_3 نشان دهنده طول مارکوف اتریوم به بیتکوین است، $l_3 \leq 1$.



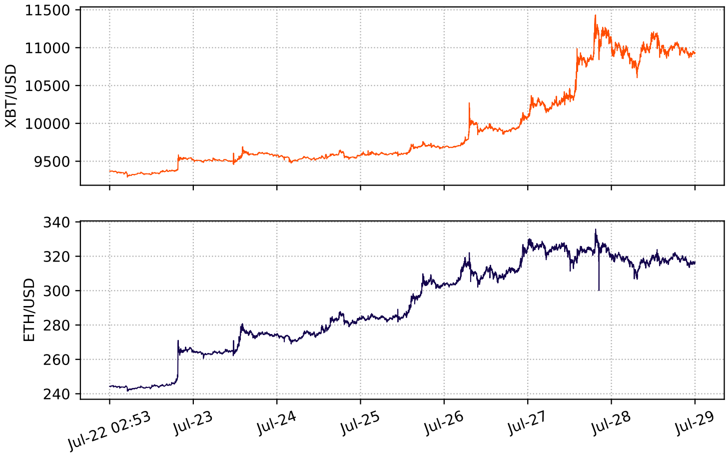
(د) در این شکل S_4 نشان دهنده طول مارکوف اتریوم به خودش است، $l_4 = 5$.

شکل ۶.۴: نمودار آزمون چپمن-کولموگروف برای بیتکوین و اتریوم در بازه شماره ۲.

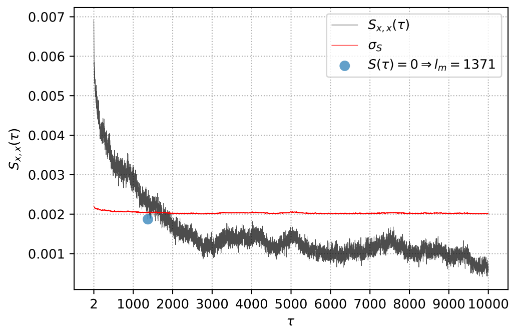
محاسبه طول مارکوف در بازه شماره ۳

بازه شماره ۳ از تاریخ ۲۲ جولای ۲۰۲۰ ساعت ۲:۵۳:۲۰ تا ۲۹ جولای ۲۰۲۰ است. مقدار انحراف معیار بیتکوین در این بازه برابر 8.93×10^{-5} و انحراف معیار اتریوم برابر با 1.69×10^{-4} است. نوسان قیمت بیتکوین و اتریوم در این بازه در شکل ۷.۴ آورده شده است. این بازه به این دلیل انتخاب شده است که انحراف معیار بیتکوین در این بازه کمتر از انحراف معیار سه ماهه بیتکوین است و انحراف معیار اتریوم بیشتر از انحراف معیار سه ماهه است.

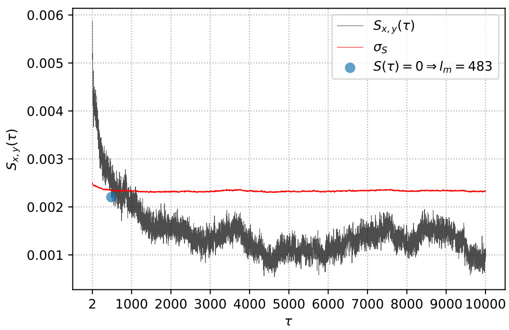
در شکل ۸.۴ نتیجه آزمون چپمن-کولموگروف برای بازه شماره ۳ رسم شده است. با توجه به شکل‌های ۸.۴آ و ۸.۴ب طول مارکوف بیتکوین نسبت به خودش و اتریوم یعنی l_1 و l_2 به ترتیب 1371 و 483 ثانیه به دست آمده‌اند. همچنین با توجه به شکل‌های ۸.۴ج و ۸.۴د طول مارکوف اتریوم نسبت به بیتکوین $l_3 = 597$ ثانیه و طول مارکوف اتریوم نسبت به خودش $l_4 = 2012$ ثانیه به دست آمده‌اند.



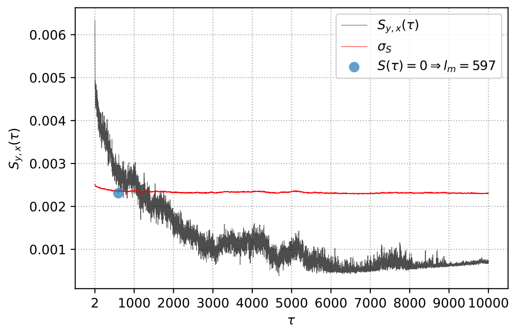
شکل ۷.۴: قیمت قراردادهای دائمی بیتکوین و اتریوم بر حسب دلار آمریکا در بازه شماره ۳.



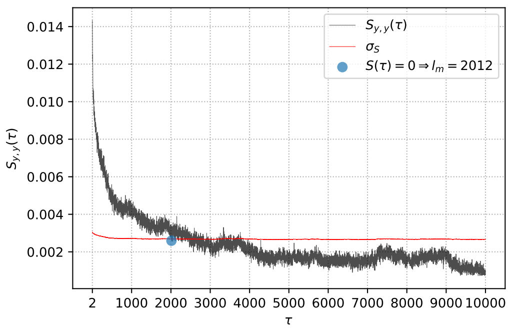
(آ) $S_1(\tau), l_1 = 1371$.



(ب) $S_2(\tau), l_2 = 483$.



(ج) $S_3(\tau), l_3 = 597$.



(د) $S_4(\tau), l_4 = 2012$.

شکل ۸.۴: نمودار آزمون چپمن-کولموگروف برای بیتکوین و اتریوم در بازه شماره ۳.

فصل ۵

نتیجه‌گیری

در فصل ۲ روشی به نام آزمون چپمن-کولموگروف معرفی شد که برای تخمین حافظه یک فرایند تصادفی یا همان طول مارکوف استفاده می‌شود و دیدیم که این روش به خوبی حافظه سری‌های زمانی تولید شده با مدل خودبرگشت را تخمین می‌زند. در ادامه در فصل ۳ آزمون چپمن-کولموگروف را برای تخمین حافظه و تخمین وابستگی دو فرایند غیرمارکوف وابسته تعمیم دادیم. به منظور آزمایش این روش به مدلی برای تولید سری‌های زمانی وابسته با طول‌های مارکوف دلخواه نیاز بود که برای این کار با تعمیم مدل خودبرگشت و اضافه کردن جملاتی برای وابسته کردن دو سری زمانی به یکدیگر موفق شدیم سری‌های زمانی وابسته با طول‌های مارکوف دلخواه تولید کنیم. در ادامه نشان دادیم که تعمیم آزمون چپمن-کولموگروف برای فرایندهای غیرمارکوف وابسته به خوبی طول مارکوف سری‌های زمانی وابسته تولید شده با استفاده از مدل ارائه شده را تخمین می‌زند همچنین وابستگی این دو سری زمانی به یکدیگر را به خوبی نشان می‌دهد.

آزمون چپمن-کولموگروف برای دو فرانید غیرمارکوف وابسته به دو شکل بر روی سری‌های زمانی تولید شده با استفاده از مدل ارائه شده بررسی شد؛ در حالت اول تعدادی زیادی نمونه مستقل از یکدیگر تولید شد. در این حالت می‌توانیم وابستگی سری‌های زمانی به خودشان و یگدیگر را در دو زمان متفاوت بررسی کنیم. اما از آنجایی که اغلب در دنیای واقعی ممکن نیست که بیش از یک سری زمانی از فرانیدهای تصادفی داشته باشیم. بنابراین با این فرض که سری‌های زمانی مانا هستند توانستیم آزمون چپمن-کولموگروف را بر حسب فاصله زمانی یعنی τ انجام دهیم.

در ادامه در فصل ۴ با استفاده از آزمون چپمن-کولموگروف تلاش کردیم تا حافظه و وابستگی دو رمزارز بیتکوین

و اتریوم را بررسی کنیم. برای این کار از قیمت قراردادهای دائمی بیتکوین و اتریوم در صرافی بیتمکس که بیشترین حجم معاملات را در میان صرافی‌های دیگر دارد استفاده کردیم. با توجه به اینکه رفتار بازارهای مالی می‌تواند وابسته به زمان باشد بنابراین محاسبه طول مارکوف را در پنجره‌های هفتگی انجام دادیم که در این پایان‌نامه نتایج سه تا از این پنجره‌ها آورده شده‌اند. با توجه به ۴.۴ یعنی جایی که نوسان قیمت بیتکوین در بیشترین مقدار خود در این بازه سه ماهه بوده می‌بینیم که طول مارکوف بیتکوین و اتریوم طول‌های نسبتا بزرگ و از مرتبه چند دقیقه هستند. اما از طرفی در بازه شماره ۲ یعنی شکل ۶.۴ یعنی جایی که نوسان قیمت در کمترین حالت خود بوده است می‌بینیم که طول مارکوف خیلی کوتاه است. این نتایج می‌تواند به این معنی باشد که اتفاقات بزرگ و مهم در این بازار حافظه‌دار و وابسته‌اند و در مواقعی که اتفاق مهمی در بازار رخ نمی‌دهد رفتار این دو رمزارز تقریبا مستقل و بدون حافظه است. اما از طرف دیگر باید توجه داشت که هرچند طول مارکوف با بزرگ شدن نوسان افزایش می‌یابد اما این مساله لزوما به این معنی نیست که هرچه بازار نوسان بزرگ‌تری داشته باشد طول مارکوف بیشتری هم خواهد داشت. نتایج شکل ۸.۴ می‌تواند نشان‌دهنده این موضوع باشد. طول مارکوف در این بازه از دو بازه دیگر بیشتر است در حالی که نوسان بیتکوین از بازه شماره ۱ کمتر است.

همانطور که دیدیم آزمون چپمن–کولموگروف تعمیم یافته با دقت خوبی طول مارکوف دو فرایند غیرمارکوف وابسته را تخمین می‌زند. از آنجایی که در بازارهای مالی، بسته به نوع بازار ممکن است سهام یا جفت ارزهای زیادی در حال معامله باشند بنابراین شاید بتوانیم با بررسی هر جفت سهام یا ارز با آزمون چپمن–کولموگروف تعمیم یافته دید کلی نسبت به ساختار آن بازار به دست آورد و مشاهده تغییرات این ساختار در طول زمان ممکن است اطلاعات جدیدی در مورد دینامیک بازار در اختیار ما قرار دهد و بینش ما را نسبت به بازار بهبود بدهد.

کتاب‌نامه

[1] Paul, Wolfgang and Baschnagel, Jörg. *Stochastic processes: from physics to finance*. Springer, Heidelberg, 2. ed ed. , 2013. OCLC: 858033687.

[2] Kampen, N. G. van. *Stochastic processes in physics and chemistry*. Elsevier, Amsterdam; London, 2007. OCLC: 1162021053.

[3] Bressloff, Paul C. *Stochastic Processes in Cell Biology*. no. 41 in Interdisciplinary Applied Mathematics. Springer International Publishing : Imprint: Springer, Cham, 1st ed. 2014 ed. , 2014.

[4] Dougherty, Edward R. *Random processes for image and signal processing*. SPI/IEEE series on imaging science & engineering. SPIE Optical Engineering Press ; Institute of Electrical and Electronics Engineers, Bellingham, Wash. : New York, 1999.

[5] Steele, J. Michael. *Stochastic calculus and financial applications*. no. 45 in Applications of mathematics. Springer, New York, 2001.

[6] Shreve, Steven E. and Shreve, Steven E. *Continuous-time models*. no. Steven E. Shreve ; 2 in Stochastic calculus for finance. Springer, New York, NY, corr. 8. print ed. , 2008. OCLC: 551922793.

[7] Islam, I. and Verick, S. The Great Recession of 2008–09: Causes, Consequences and Policy Responses. in Islam, Iyanatul and Verick, Sher, eds. , *From the Great Recession to Labour Market Recovery*, pp. 19–52. Palgrave Macmillan UK, London, 2011.

[8] Cao, W., Demazeau, Y., Cao, L., and Zhu, W. Financial crisis and global market couplings. in *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–10, 2015.

[9] Shumway, Robert. *Time series analysis and its applications : with R examples*. Springer, New York, 2011.

[10] Gandhi, Vaibhav. Chapter 2 - interfacing brain and machine. in Gandhi, Vaibhav, ed. , *Brain-Computer Interfacing for Assistive Robotics*, pp. 7 – 63. Academic Press, San Diego, 2015.

[11] Mills, Terence C. *Time series techniques for economists*. Cambridge University Press, Cambridge [England] ; New York, 1990.

[12] Asteriou, Dimitrios and Hall, S. G. *Applied econometrics*. Palgrave Macmillan, Basingstoke [England] ; New York, 2nd ed ed. , 2011. OCLC: ocn682891193.

[13] Box, George E. P., Jenkins, Gwilym M., Reinsel, Gregory C., and Ljung, Greta M. *Time series analysis: forecasting and control*. Wiley series in probability and statistics. John Wiley & Sons, Inc, Hoboken, New Jersey, fifth edition ed. , 2016.

[14] Stroe-Kunold, Esther, Stadnytska, Tetiana, Werner, Joachim, and Braun, Simone. Estimating long-range dependence in time series: An evaluation of estimators implemented in R. *Behavior Research Methods*, 41(3):909–923, August 2009.

[15] Ghasemi, Fatemeh, Sahimi, Muhammad, Peinke, J., Friedrich, R., Jafari, G. Reza, and Tabar, M. Reza Rahimi. Markov analysis and Kramers-Moyal expansion of nonstationary

stochastic processes with application to the fluctuations in the oil price. *Physical Review E*, 75(6):060102, June 2007.

[16] Waechter, M., Riess, F., Schimmel, Th., Wendt, U., and Peinke, J. Stochastic analysis of different rough surfaces. *The European Physical Journal B - Condensed Matter and Complex Systems*, 41(2):259–277, September 2004.

[17] Fazeli, S M, Shirazi, A H, and Jafari, G R. Probing rough surfaces: Markovian versus non-Markovian processes. *New Journal of Physics*, 10(8):083020, August 2008.

[18] Risken, Hannes. Fokker-Planck Equation. in Risken, Hannes, ed. , *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer Series in Synergetics, pp. 63–95. Springer, Berlin, Heidelberg, 1984.

[19] Friedrich, Rudolf, Peinke, Joachim, Sahimi, Muhammad, and Reza Rahimi Tabar, M. Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87–162, September 2011.

[20] Friedrich, R. and Peinke, J. Description of a Turbulent Cascade by a Fokker-Planck Equation. *Physical Review Letters*, 78(5):863–866, February 1997.

[21] Ross, Sheldon M. *A first course in probability*. Pearson, Boston, ninth edition ed. , 2014.

[22] Papoulis, Athanasios and Pillai, S. Unnikrishna. *Probability, random variables, and stochastic processes*. McGraw-Hill, Boston, Mass., 4. ed., internat. ed., nachdr ed. , 2009. OCLC: 255904469.

[23] Taylor, Howard M. and Karlin, Samuel. *An introduction to stochastic modeling*. Academic Press, San Diego, 3rd ed ed. , 1998.

[24] Wikipedia contributors. Skewness — Wikipedia, the free encyclopedia, 2020.

[25] Walck, C. Hand-book on statistical distributions for experimentalists. 1996.

[26] Garcia-Palacios, J. L. Introduction to the theory of stochastic processes and Brownian motion problems. *arXiv e-prints*, pp. cond–mat/0701242, January 2007.

[27] Gagniuc, Paul A. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, Hoboken, NJ, 2017.

[28] Brémaud, Pierre. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. no. 31 in Texts in applied mathematics. Springer, New York, 1999.

[29] Kramers, H.A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, April 1940.

[30] Moyal, J. E. Stochastic Processes and Statistical Physics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(2):150–210, 1949.

[31] Pawula, R. Generalizations and extensions of the Fokker- Planck-Kolmogorov equations. *IEEE Transactions on Information Theory*, 13(1):33–41, January 1967.

[32] Lemons, Don S. and Gythiel, Anthony. Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) **146** , 530–533 (1908)]. *American Journal of Physics*, 65(11):1079–1081, November 1997.

[33] Siefert, Malte and Peinke, J. Joint multi-scale statistics of longitudinal and transversal increments in small-scale wake turbulence. *Journal of Turbulence*, (7):N50, 2006.

[34] Kimiagar, S., Sadegh Movahed, M., Khorram, S., and Reza Rahimi Tabar, M. Markov Properties of Electrical Discharge Current Fluctuations in Plasma. *Journal of Statistical Physics*, 143(1):148–167, April 2011.

[35] Kimiagar, S, Jafari, G R, and Reza Rahimi Tabar, M. Markov analysis and Kramers–Moyal expansion of the ballistic deposition and restricted solid-on-solid models. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(02):P02010, February 2008.

[36] Friedrich, R, Zeller, J, and Peinke, J. A note on three-point statistics of velocity increments in turbulence. *Europhysics Letters (EPL)*, 41(2):153–158, January 1998.

[37] Wilcoxon, Frank. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80, December 1945.

[38] Renner, Christoph, Peinke, J., and Friedrich, R. Experimental indications for Markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, 433:383–409, April 2001.

[39] Tutkun, Murat and Mydlarski, Laurent. Markovian properties of passive scalar increments in grid-generated turbulence. *New Journal of Physics*, 6(1):49, 2004.

[40] Prinz, Jan-Hendrik, Wu, Hao, Sarich, Marco, Keller, Bettina, Senne, Martin, Held, Martin, Chodera, John D., Schütte, Christof, and Noé, Frank. Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics*, 134(17):174105, May 2011.

[41] BitMEX. Historical extracts of public bitmex api, public.bitmex.com, 2020.

[42] Noé, Frank, Keller, Bettina, and Prinz, Jan-Hendrik. Lecture notes on stochastic processes. 2013.

[43] Norris, J. R. *Markov chains*. Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, Cambridge, UK ; New York, 1st pbk. ed ed. , 1998.

Abstract

There are many systems in the real world showing stochastic behaviors. We use stochastic processes to model such behaviors. Since no system evolves entirely independent and interacts with other systems, their individual study could lead to vague results. In addition to such interdependency, a system may also depend on its own past, or in other words, it has memory. In the context of stochastic processes a process with memory is called a non-Markov process. Financial markets are an example of complex systems whose behavior can be modeled using stochastic processes with memory; they interact with each other in a sense that they are influenced by each other past and also themselves. Given the role of complex systems such as financial markets in human life, it is necessary to have methods for estimating memory and the dependency of stochastic processes. One of the most common methods for estimating the memory of a non-Markov process (Markov length) is the Chapman-Kolmogorov test. This dissertation aims to generalize the Chapman-Kolmogorov test into two dependent non-Markov processes. To show the efficiency in estimating the Markov length for two dependent non-Markov processes, we need a model that generates two dependent time series with the desired Markov lengths. The results show that the generalized Chapman-Kolmogorov test accurately estimates Markov lengths in two dependent non-Markov processes. In the end, we analyzed Bitcoin and Ethereum, which are cryptocurrencies. Results show that significant moves have memory, but small ones are independent of each other.

Keywords:

non-Markov dependent processes, Chapman-Kolmogorov test, Memory, Markov length, Autoregressive model, Cryptocurrency, Bitcoin, Ethereum

Shahid Beheshti University

Department Physics

A dissertation submitted for the degree of Master of Science in Physics

(Complex Systems and Nonlinear Dynamic)

Modeling and memory estimation for dependent non-Markov processes

Supervisers:

Dr. A. Hosseini

Dr. G. Jafari

Adviser:

Dr. T. Jamali

By: Hossein Motahari

September 2020