بسم الله الرحمن الرحيم

دانشگاه تربیت معلم

دانشکده علوم ریاضی و کامپیوتر

پایان نامه کارشناسی ارشد

آمارریاضی

عنوان

آشکارسازی نقاط تغییر در زنجیر مارکوف

تدوین

عبدالصالح توغدری

استاد راهنما

دکتر علی اکبر رحیم زاده ثانی

تیر 1388

چکیده

زنجیرهای مارکوف مدل انعطاف پذیری را برای متغیرهای تصادفی وابسته، با کاربردهای در زمینه‌های نظیر فیزیک، علوم طبیعی و اقتصاد فراهم می‌کنند. در یک مطالعه‌ی کاربردی از یک زنجیر مارکوف، موضوع مورد علاقه ممکن است تشخیص تغییرات ماتریس احتمال تغییر وضعیت، در مشاهده‌ی تحققی از فرایند باشد. اگر چنین تغییراتی رخ دهند آنگاه، برآورد مکانهایی که تغییرات در آنجا رخ می‌دهند و برآورد ماتریس احتمالهای تغییر وضعیت قبل و بعد از تغییرات موضوع مورد علاقه خواهد بود. وقتی تعداد نقاط تغییر معلوم است. نظریه‌ی درستنمایی استاندارد برای آشکار کردن مکان آنها به کار می‌رود، و از بوت استرپ برای محاسبه‌ی P-مقدار استفاده می‌کنیم. وقتی تعداد نقاط تغییر نامعلوم است، اندازه‌های AIC و BIC برای انتخاب مدل به کار می‌روند. روشهای پیشنهاد شده به طور تجربی مطالعه شده‌اند و همچنین برای مجموعه‌ای از مثالها به صورت کاربردی اعمال شده‌اند.

واژگان کلیدی: AIC ، BIC ، بوت استرپ، نقطه-تغییر، زنجیر مارکوف، درستنمایی

رده‌بندی موضوعی، آمارریاضی(2006): 62M02 ، 60J10

فهرست مطالب

پیشگفتار

در این پایان‌نامه هدف معرفی نحوه‌ی آشکار کردن نقاط تغییر در زنجیر مارکوف زمان گسسته به کمک نظریه درستنمایی مجانبی است. مزایای آن این است که می‌توانیم جاهایی را که در آنها تغییراتی در فرایند مارکوف رخ می‌دهند آشکار کنیم. این موضوع توسط آلن م.پولانسکی[3] معرفی شد و محور اصلی آن مقاله‌ی [3] است. مدل زنجیر مارکوفی که می‌خواهیم در این پایان نامه کار - کنیم، ماتریس احتمال تغییر وضعیت را برای تغییر در طی تحقق مشاهده شده مورد بررسی قرار می- دهد. بنابراین، در این مدل دنباله‌ی ماتریس‌های احتمال تغییر وضعیت $\mathrm{T_o}$، … ،T_k و اعداد صحیح مثبت $\mathrm{o} = \psi_o < \psi_1 < \cdots < \psi_{k+1} = n$ که در آن $\mathrm{P_j} = \mathrm{T_i}$ برای $\mathrm{j} = \psi_i, \dots, (\psi_{i+1} - 1)$ و $\mathrm{i} = 1, \dots, \mathrm{k}$ وجود دارند که، نقاط تغییر ψ_1، … ،ψ_k معلوم فرض می‌شوند. اگر زمان مداخله‌ها معلوم باشند یعنی نقاط تغییر ψ_1، … ،ψ_k معلوم باشند، در چنین موقعیتی برآورد $\mathrm{T_1}$، … ،T_k همانند آزمون فرض صفر به فرم $\mathrm{H_o}: \mathrm{T_o} = \cdots = \mathrm{T_k}$ در مقابل، فرض مقابل برای حداقل یک $\mathrm{i} \neq \mathrm{j}$ ، $\mathrm{H_1}: \mathrm{T_i} \neq \mathrm{T_j}$ مفید است. این امر محقق را مجاز می‌کند که تحقیق کند آیا نتایج مداخله‌ها در رفتار فرایند تغییرات معنی‌داری می‌دهند یا خیر. در موارد دیگر ، ممکن است نقاط تغییر نامعلوم باشند در چنین مواردی ، ابتدا باید تعداد و سپس مکان نقاط تغییر را برآورد کنیم. روشهای بیزی برای حل مسئله‌ی نقطه تغییر تنها به وسیله‌ی (سیلوا ومالتسو[28]) وسیلوا[36] ویاکر[43] تحقیق شده است .به طور کلی این روشها روی استراتژی توقفهای بهینه برای کشف از یک تغییر تنها در ماتریس احتمال تغییر وضعیت برای مشاهدات در طول زمان از یک فرایند در حال اجرا متمرکز شده است. تاکید این پایان نامه کشف نقاط تغییر روی اندازه‌ی نمونه‌ی ثابت است. یک آزمون عمومی برای همگنی برای دنباله‌های مارکوف زمان گسسته به این صورت است که دنباله‌ای از پارامترهای تحت فرض صفر یک دنباله از متغیرهای تصادفی مستقل به طور یکنواخت گسسته توزیع شده است این نوع آزمون‌ها به وسیله‌ی راجارشی و راماندم [32]مطالعه شده است. در این پایان نامه فرض می‌شود پارامترها قابل برآورد هستند. اگر نمونه‌های چندگانه قابل دسترسی

هستند آنگاه روش توسعه داده شده توسط اندرسن وگودمن [1] و مادانسکی [27] می‌تواند روش کاملی برای کشف کردن نقاط تغییر وقتی حداقل یک نقطه تغییر رخ دهد، باشد. روشهای به کار رفته در پایان نامه مخصوصا روی مسئله‌ای از یک تحقق مشاهده شده تنها و سرانجام مفروضاتی از مسئله چند نقطه‌ی تغییر در یک دنباله‌ی مشاهده شده تمرکز می‌کند. روشهای این پایان نامه بر پایه‌ی نظریه درستنمایی مجانبی است. بازبینی کلی از نتایج مجانبی برای مسئله‌های نقطه تغییر به وسیله-ی هوروات وکسورگو[12] ارائه شده است.

برای ارائه این مفاهیم از مقالات زیر به عنوان مقاله اصلی استفاده شده است.

[3] Alan M. Polansky. 2006. Detecting change-points in Markov chains. Division of Statistics, Northern Illinois University De Kalb, IL ,USA, 6013-6026.

[8] Billingsley,P., 1961 a. Statistical method in Markov chains. Ann. Math. Statist. 32,12-40.

از مقاله‌های زیر نیز به عنوان مقالات فرعی استفاده شده است.

[16] Efron, B., 1979. The bootstrap another look at the jackknife. Ann. Statist. 7, 1-26.

[2] Akaike, H., 1974. A new look at statistical model identification. IEEE Trans. Automat. Control 19, 716-723.

[5] Athreya, K.B., Fuh, C.D., 1992. Bootstrapping Markov chains. In: lePage, R., Billard, L. (Eds)., Exploring the Limits of Bootstrapping. Wiley, New York, pp. 49-64.

در فصل اول کلیات و تعاریف مورد نیاز برای موضوع را می‌آوریم، و در فصل دوم قضایا و روشهای مورد استفاده در آشکار کردن نقاط تغییر را می‌آوریم و مقدمه‌ای برای نقطه تغییر بیان می‌کنیم، نظریه درستنمایی مجانبی و خی دو را بیان و قضایای مربوط به آن را اثبات می‌کنیم. در فصل سوم ، به موضوع آشکار کردن نقاط تغییر به کمک نظریه درستنمایی مجانبی می‌پردازیم، ابتدا روشها و

تکنیکهای مورد نیاز را بیان می‌کنیم و سپس به صورت تجربی آنها را بررسی می‌کنیم، در آخر کار تعمیم روشها را ارائه کرده و سپس چند مثال کاربردی ارائه می‌دهیم و در پایان نتیجه گیری را بیان می‌کنیم.

فصل اول

کلیات و تعاریف

فصل اول

کلیات وتعاریف

1-1. فرایندهای تصادفی

فرض می‌کنیم T مجموعه‌ای دلخواه و به ازای هر $t \in T$، X_{t} متغیری تصادفی باشد. فرض می‌کنیم $E \subset R$ مجموعه‌ی ثابتی باشد ومقادیر متغیرهای تصادفی X_{t} در داخل این مجموعه باشد. در این صورت تعریف زیر را داریم[49]:

1-1-1. **تعریف**: مجموعه‌ی $X = \{X_{\mathrm{t}}: t \in T\}$ را فرایند تصادفی با مجموعه‌ی اندیسگذار T و فضای وضعیت E می‌گوییم، اگر $A \subset E$ (یا $x \in E$) و $X_{\mathrm{t}} \in A$ (یا $X_{\mathrm{t}} = x$) می‌گوییم فرایند در زمان (یا مرحله‌ی t) در مجموعه‌ی A (یا در وضعیت x) قرار دارد. اگر T مجموعه‌ای شمارا باشد، فرایند را گسسته زمان(یا زمان گسسته) و اگر مجموعه‌هایی به صورت $(\infty\,,.)$ و$(-\infty,\infty)$ باشد آن را پیوسته زمان (یا زمان پیوسته) می‌گوییم. برای هر ω از فضای نمونه‌ای مجموعه‌ی $\{X_{\mathrm{t}}(\omega): t \in T\}$ را که زیر مجموعه‌ای از E است، تحقق یا مسیر نمونه‌ای فرایند می‌گوییم.

2-1-1. **تعریف**: **(الف)**. در فرایند $\{X_{\mathrm{t}}: t \geq o\}$ تفاضل $X_{\mathrm{t}} - X_{\mathrm{s}}$ و $\mathrm{t} > s$ ، را نمو فرایند در بازه‌ی (s,t) می‌گوییم. اگر ساختمان نموها معلوم باشند می‌توان از آنها در ساختن خود فرایند و یا پاسخ دادن به مسئله‌های اساسی فرایند استفاده کرد. مثلا در زمانهای 5 و10 به ترتیب در وضعیت-های 4 و 15 باشد، آنگاه نمو آن در فاصله‌ی (5،10) برابر 11= 4 - 15، است.

(ب). فرایند $\{X_{\mathrm{t}}: t \geq o\}$ را با نموهای مانا می‌گوییم، اگر به ازای هر $\mathrm{t} > s$ توزیع نمو $X_{\mathrm{t}} - X_{\mathrm{s}}$ فقط به تفاضل t و s بستگی داشته باشد. یعنی نموهای فرایند در بازه‌هایی با طولهای برابر از نظر احتمالاتی دارای یک ساختمان باشند. لذا در این گونه فرایندها مبدا زمان تاثیری ندارد.

(پ). فرایند $\{X_{\mathrm{t}}: t \geq o\}$ را مانای اکید می‌گوییم هرگاه به ازای هر $n \geq 1$ و هر t_1 ، t_2 ، … ، t_n و هر $h \geq o$ توزیع توام X_{t_1}، … ، $\mathrm{X}_{\mathrm{t_n}}$ همان توزیع توام متغیرهای تصادفی $X_{t_{1+h}}$، … ، $\mathrm{X}_{\mathrm{t}_{\mathrm{n}+h}}$ باشد. با فرض $n = 1$، نتیجه می‌شود که در هر فرایند مانای اکید به ازای هرt و s ، X_{t} و X_{s} هم توزیع‌اند $(h = \mathrm{t} - s)$ و لذا توزیع X_{t} به t بستگی ندارد. یعنی تمام متغیرهای تصادفی تشکیل دهنده‌ی فرایند، همتوزیع‌اند. از اینجا نتیجه می‌شود که $E(X_{\mathrm{t}})$ به t بستگی ندارد و مقدار ثابتی است. با

فرض $n=2$ نتیجه می‌شود که در فرایند مانای اکید به ازای هر t و s، توزیع توام X_{t} و X_{s} وتوزیع توام X_o و $X_{\mathrm{t}-s}$ باهم برابراند $(h=s)$، ولذا توزیع توام X_{t} و X_{s} فقط بستگی به تفاضل t و s دارد. بنابراین هرگونه محاسبه مبتنی بر توزیع توام X_{t} و X_{s} به تفاضل t و s بستگی دارد. از جمله محاسباتی که براساس توزیع انجام می‌گیرد. تابع کوواریانس است. بنابراین $C(X_{\mathrm{t}},X_{\mathrm{s}})$ تابعی از $\mathrm{t}-s$ است. چون در این گونه فرایندها $E(X_{\mathrm{t}})$ و $E(X_{\mathrm{s}})$ ثابت‌اند و همچنین

$$C(X_{\mathrm{t}},X_{\mathrm{s}})=E(X_{\mathrm{s}}X_{\mathrm{t}})-E(X_{\mathrm{s}})\,E(X_{\mathrm{t}}).$$

نتیجه می‌شود که در فرایندهای مانای اکید، $E(X_sX_t)$ به $t-s$ بستگی دارد.

(ت). فرایند $\{X_{\mathrm{t}}:t\geq o\}$ را فرایند مانای وسیع می‌گوییم هرگاه $E(X_{\mathrm{t}})$ ثابت باشد، و $E(X_sX_t)$ فقط به تفاضل t و s بستگی داشته باشد.

3-1-1. **فرایند مارکوف** :

فرض می‌کنیم $\{X(n);n=0،1،2،\ldots\}$ یک فرایند تصادفی زمان گسسته و فضای وضعیت شمارای l باشد. گوییم این فرایند یک فرایند مارکوف است اگر به ازای هر $n\geq 1$, هر x_o،...،$x_{\mathrm{n}-1}$، x، y از وضعیت‌ها، برابری

$$P\left(X_{n+1}=y|X_n=x،\ldots،X_1=x_1،\mathrm{X_o}=x_{\mathrm{o}}\right)=P\left(X_{n+1}=y|X_n=x\right).$$

برقرار باشد. یعنی فقط اطلاع از وضعیت فرایند در مرحله‌ی n برای تعیین توزیع وضعیت فرایند در مرحله‌ی $n+1$ کفایت می‌کند و اطلاعات قبل از آن موثر نخواهد بود. این ویژگی را ویژگی مارکوفی می‌گویند. احتمال شرطی $P(X_{n+1}=y|X_n=x)$ را احتمال تغییر وضعیت یک مرحله (از x در مرحله‌ی nام به y در مرحله‌ی $n+1$ام) می‌نامیم. اگر این احتمال به n بستگی نداشته باشد، چنین فرایند مارکوفی را یک فرایند مارکوف همگن می‌گویند.

حال، اگر $\{X(t);t\geq o\}$ فرایندی تصادفی زمان پیوسته باشد که مقادیرش را در مجموعه‌ی اعداد حقیقی اختیار می‌کند، می‌گوییم فرایند $\{X(t);t\geq o\}$ فرایند مارکوف زمان پیوسته است اگر به ازای مقادیر دلخواه $\mathrm{t_o}<t_1<\cdots<t_n=s$ و اعداد حقیقی $x_o،\ldots،x_n=x$،

داشته باشیم،

$$P\left\{X(t+s) \leq y \,\middle|\, X(s)=x، X\left(t_{n-1}\right)=x_{n-1}، \ldots ،X(t_o)=x_o\right\} =$$
$$P\{X(t+s) \leq y|X(s)=x\}.$$

به عبارت دیگر، فرایند مارکوف زمان پیوسته، فرایندی است، تصادفی با این ویژگی مارکوفی که توزیع شرطی وضعیت آینده در زمان $t+s$ با معلوم بودن وضعیت فعلی در s و تمام وضعیت‌های گذشته، فقط به وضعیت فعلی بستگی دارد و از گذشته مستقل است، به علاوه اگر

$$P\{X(t+s) \leq y|X(s)=x\}.$$

مستقل از s باشد، آنگاه گوییم فرایند مارکوف زمان پیوسته دارای احتمالهای تغییر وضعیت مانا یا همگن است. در حالت همگن $P(X_n=y|X_o=x)$ را به صورت $P_{\mathrm{xy}}^{(\mathrm{n})}$ نشان می‌دهیم. $P_{\mathrm{xy}}=P_{\mathrm{xy}}^{(1)}$ را احتمال تغییر وضعیت از وضعیت x به وضعیت y در یک واحد زمان می‌نامیم.

1-2. **وضعیتهای گذرا و بازگشتی**:

براساس رفتار احتمالهای تغییر وضعیت ، حالتهای زنجیر مارکوف را به وضعیتهای گذرا و بازگشتی تقسیم می‌کنیم[49].

1-2-1. **تعریف**: گوییم وضعیت y در دسترس وضعیت x است، هرگاه به ازای عددی مانند n ، $n \geq o$ و $P_{\mathrm{xy}}^{(\mathrm{n})} > \mathrm{o}$، به عبارت دیگر بتوانیم در تعدادی متناهی از مراحل از x به y برویم. اگر y در دسترس x باشد، از نماد $x \to y$ استفاده می‌کنیم. اگر x و y در دسترس یکدیگر باشند آنگاه می-نویسیم. $x \leftrightarrow y$. اگر فرایند در وضعیت x قرار دارد زمان اولین رسیدن به y را با $\mathrm{T_y}$ نشان می-دهیم بنابراین $\mathrm{T_y}=min\{n: X_n=y|X_o=x\}$. پیشامد رفتن از x به y در تعدادی متناهی از مراحل را می‌توان به پیشامدهای ناسازگار $\{\mathrm{T_y}=n\}$ ، $n=1$،... تجزیه کرد، بنابراین

$$P_{\mathrm{x}}\{\text{رسیدن به} y \text{ درتعداد متناهی ازمراحل}\} = \mathrm{P_x}\{\mathrm{T_y}<\infty\} = \mathrm{P_x}\{\cup_{n=1}^{\infty}\{\mathrm{T_y}=n\}\} =$$
$$\textstyle\sum_{\mathrm{n}=1}^{\infty} \mathrm{P_x}\{\mathrm{T_y}=n\}.$$

با قرار دادن $\mathrm{P_x}\{\mathrm{T_y}=n\}=f_{\mathrm{xy}}^{(\mathrm{n})}$ ،

$$P_{\mathrm{x}}\{\text{رسیدن به y درتعداد متناهی ازمراحل}\} = f_{\mathrm{xy}}^{(\mathrm{n})}.$$

رابطه زیر بدست می‌آید: $f_{\mathrm{xy}}^{(\mathrm{n})} = \mathrm{P}_{\mathrm{x}}\{\mathrm{T}_{\mathrm{y}} < \infty\} = \sum_{\mathrm{n}=1}^{\infty} f_{\mathrm{xy}}^{(\mathrm{n})}$. لازم است تمایز بین $f_{\mathrm{xy}}^{(\mathrm{n})}$ و $P_{\mathrm{xy}}^{(\mathrm{n})}$ تاکید شود. $P_{\mathrm{xy}}^{(\mathrm{n})}$ احتمال شروع از x و بودن در y در مرحله n ام است. در این احتمال وضعیتهای که فرایند در مرحله‌های بین 1 و n به خود می‌گیرد، دخالتی ندارند، در صورتی که در $f_{\mathrm{xy}}^{(\mathrm{n})}$ درست است که احتمال شروع از x و بودن در y در مرحله‌ی n ام مورد توجه است ، اما در بین مراحل 1تا n فرایند نباید در حالت y قرار گرفته باشد یا به عبارت دیگر

$$f_{\mathrm{xy}}^{(\mathrm{n})} = P(X_n = y; X_{n-1} \neq y, \ldots, X_1 \neq y | X_o = x).$$
$$P_{\mathrm{xy}}^{(\mathrm{n})} = P(X_n = y | X_o = x) .$$

از مقایسه دو احتمال بالا معلوم می‌شود که همواره

$$f_{\mathrm{xy}}^{(\mathrm{n})} \leq P_{\mathrm{xy}}^{(\mathrm{n})} .$$

1-2-2. **تعریف**: وضعیت x را گذرا می‌گوییم، هرگاه $f_{\mathrm{xx}} < 1$ و آن را بازگشتی گوییم هرگاه $f_{\mathrm{xx}} = 1$. به عبارت دیگر اگر x گذرا باشد با شروع از x با احتمال مثبت $1\text{-}f_{\mathrm{xx}}$ ممکن است هیچ گاه به x باز نگردیم، اما اگر x بازگشتی باشد با شروع از آن بالاخره در یک زمان متناهی حتما به آن باز خواهیم گشت. اگر زنجیر بازگشتی مثبت باشد (یعنی وقتی از یک وضعیت زنجیر خارج می-شویم حتما در یک زمان متناهی و با احتمال مثبت به آن بر خواهیم گشت.) زنجیر را ارگودیک می-گوییم.

1-3. تابع درستنمایی

بفرض X متغیر تصادفی با چگالی $f(x;.)$ و با پارامتر $\theta\epsilon\Theta$ و $x \in R$ باشد، اگر X_1، ... ،X_n نمونه‌ی تصادفی از چگالی $f(x;\theta)$ باشد، آنگاه تابع درستنمایی برابر

$$L\left(\theta; x_1, x_2, \ldots, x_n\right) = f(x_1;\theta). f(x_2;\theta) \ldots f(x_n;\theta)$$ و $\theta\epsilon\Theta$ ، $i = 1, \ldots, n$ و

$$\mathrm{x_i} \in R$$

می‌باشد که در آن نمادگذاری برای تداعی این که تابع درستنمایی تابعی از θ است، نماد $L(\theta; x_1, \dots, x_n)$ یا $L(.; x_1, \dots, x_n)$ را برای تابع درستنمایی به کار می بریم. تابع درستنمایی $L(\theta; x_1, \dots, x_n)$، درستنمایی‌ای را که متغیرهای تصادفی مقدار ویژه‌ی x_1، ... ،x_{n} را اختیار کنند، می‌دهد. بنابراین برای متغیرهای تصادفی گسسته، برابر با احتمال وقوع آنها می‌باشد. برای یک لحظه فرض کنیم θ معلوم است و مقدار آن را با θ_o نشان دهیم. مقدار ویژه‌ی متغیرهای تصادفی که «بیشتر محتمل است رخ دهد » آن مقدار x'_1، ... ،x'_{n} است، به طوری که

$$f_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n ; \theta_o).$$

ماکزیمم باشد[50].

1-3-1. برآوردگر درستنمایی ماکزیمم

فرض می‌کنیم $L(\theta) = L(\theta; x_1, \dots, x_n)$ تابع درستنمایی برای متغیرهای تصادفی X_1، ... ،$\mathrm{X_n}$ باشد. اگر $\hat{\theta}$ (که $\hat{\Theta} = \hat{\vartheta}(x_1, \dots, x_n)$ تابعی از مشاهدات $x_1, \dots, x_n$ می‌باشد) مقدار θ در Θ باشد که $L(\theta)$ را ماکزیمم می‌کند، آنگاه $\hat{\Theta} = \hat{\vartheta}(\mathrm{X}_1, \dots, X_n)$ برآوردگردرستنمایی ماکزیمم θ برای نمونه‌ی X_1، ... ،$\mathrm{X_n}$ است.

1-4. سازگاری

هنگام بررسی دنباله‌ای از برآوردگرها، به نظر می رسد که یک دنباله‌ی خوب از برآوردگرها، دنباله‌ای است که برای آنها با افزایش حجم نمونه، مقادیر برآوردگرها به مقدار واقعی پارامتر که باید برآورد شود میل کنند.

1-4-1. تعریف: سازگاری میانگین توان دوم خطا

فرض می‌کنیم $\mathrm{T}_1, \dots \mathrm{T}_n, \dots$ دنباله‌ای از برآوردگرهای $\tau(\theta)$ باشد، که درآن $\mathrm{T_n} = \mathrm{t_n}(X_1, \dots, X_n)$ براساس یک نمونه به حجم n می‌باشد. این دنباله از برآوردگرها یک دنباله از برآوردگرهای سازگار در میانگین توان دوم خطا گفته می‌شود اگر برای هر θ در Θ

$$\lim_{n\to\infty} E_\theta\left[\left(T_n-\tau(\theta)\right)^2\right]=\text{٠}.$$

1-5. **P-مقدار:آماره آزمون برای ارزیابی**

افراد مختلفی در مورد مسئله آزمون یکسان، ممکن است، ناحیه‌ی بحرانی مختلفی ازاندازه α داشته باشند. آزمایشگر اول ممکن است، فرض H_o را با استفاده از آزمونی با اندازه‌ی $\alpha=.05$ رد کند. در حالی که آزمایشگر دوم همین آزمون را با اندازه‌ی $\alpha=.01$ به کار برد. یک نمونه تصادفی از جامعه گرفته می‌شود که هم نمونه و هم آماره آزمونی که هردو آزمایشگر به کار می‌برند، یکی است، آزمایشگر اول ممکن است فرض H_o را رد کند، در حالی که آزمایشگر دوم فرض H_o را بپذیرد. این مشکل را می‌توانیم با گزارش دادن نتیجه آزمایش براساس P-مقدار، احتمال معنی داری آزمون حل کنیم[31].

1-5-1. **تعریف: P-مقدار**

P-**مقدار** کمیتی است، آماری که کمترین سطح معنی‌داری از اندازه‌ی α تعریف می‌شود که یک آزمایشگر با کاربرد آماره آزمون T روی نمونه‌ی مشاهده شده از جامعه فرض H_o را رد می‌کند. یعنی اگر مقدار بحرانی، αی تعیین شده توسط آزمایشگر از P-مقدار محاسبه شده کمتر باشد، فرض H_o رد نمی‌شود، و در غیر این صورت فرض H_o رد می‌شود.

1-6. **ماتریس**:

یک ماتریس $m\times n$ مانند A یک آرایه‌ی مستطیلی از mn عدد است که به صورت m سطر و n ستون است آرایش یافته است:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix}.$$

مولفه‌ی ij ام A که با a_{ij} نشان داده می‌شود، عددی است که در سطر i ام و ستون j ام A آمده است. ماتریس A را به صورت $A=\left(a_{ij}\right)$ نشان می‌دهیم. ماتریس‌ها را معمولا با حروف بزرگ نشان می‌دهیم، و اگر یک ماتریس $m\times n$ ، $m=n$ باشد، آنگاه A را یک ماتریس مربعی می‌نامیم.

یک ماتریس $m \times n$ که همه‌ی مولفه‌های آن صفر باشند، ماتریس صفر $m \times n$ نام دارد، و ماتریسی که فقط عناصر روی قطر اصلی آن یک باشند وعناصر غیر قطر اصلی آن همگی صفر باشند را ماتریس همانی می‌نامیم و با I نشان می‌دهیم. گوییم یک ماتریس $m \times n$ دارای اندازه‌ی $m \times n$ است. دو ماتریس $A = (a_{ij})$ و $B = (b_{ij})$ مساوی‌اند اگر (1). دارای یک اندازه باشند (2). مولفه‌های نظیر به نظیر آنها مساوی باشند[47] و [48].

1-6-1. **تعریف**: فرض می‌کنیم $A = (a_{ij})$ یک ماتریس مربع $k \times k$ باشد. اثر ماتریس A که به صورت $tr(A)$ نوشته می‌شود، مجموع درایه‌های قطری است، یعنی، $tr(A) = \sum_{i=1}^{k} a_{ii}$.

ویژگی‌های $tr(.)$: فرض می‌کنیم A و B دو ماتریس $k \times k$ و c یک اسکالر باشد، آنگاه داریم:

الف). $tr(cA) = ctr(A)$.

ب). $tr(A \pm B) = tr(A) \pm tr(B)$.

ج). $tr(AB) = tr(BA)$.

د). $tr(B^{-1}AB) = tr(A)$.

2-6-1. ماتریس مربع A را متعامد می‌گوییم هرگاه سطرهایش که به صورت بردار در نظر گرفته می-شوند، دو به دو متعامد بوده و طول واحد داشته باشند، یعنی، $A\acute{A} = I$ که در آن $\acute{A}$ را ترانهاده‌ی ماتریس A می‌گوییم یعنی اگر $A = (a_{ij})$ و $\acute{A} = (b_{ji})$ در این صورت $a_{ij} = b_{ji}$.

3-6-1. **نتیجه**: ماتریس A متعامد است اگر وتنها اگر $\acute{A} = A^{-1}$. برای یک ماتریس متعامد داریم: $\acute{A}A = A\acute{A} = I$. بنابراین ستونها نیز دو به دو متعامد بوده و دارای طول واحد هستند.

فرض می‌کنیم A ماتریس مربع $k \times k$ و I ماتریس همانی $k \times k$ باشد. در این صورت اسکالرهای λ_1، ...، λ_k را که در معادله چندجمله‌ای $|A - \lambda I| = o$ صدق می‌کنند، مقادیر ویژه (یا ریشه‌های مفسر) ماتریس A می‌نامند. معادله‌ی $|A - \lambda I| = 0$ به عنوان تابعی از λ را معادله مفسر می‌نامیم.

در اینجا، دترمینان ماتریس $k \times k$، A را با $|A|$ نشان می‌دهیم.

1-7. **فرمول ویتل**[1]: فرض می‌کنیم X_1، ...، X_{n+1} یک نمونه از فرایند مارکف مرتبه اول با احتمالهای تغییر وضعیت p_{ij} و احتمال آغازی p_{i} (احتمال هر وضعیتی که رخ می‌دهد) و اگر a_1، ...، a_{n+1} یک مجموعه‌ای به طول $n+1$ از وضعیتها باشد آنگاه احتمال اینکه X_1، ...، X_{n+1} با هم رخ دهند برابر است با:

$$P(X_1 = a_1, X_2 = a_2, \dots, X_{n+1} = a_{n+1}) =$$
$$P(X_1 = a_1)P(X_2 = a_2|X_1 = a_1) \dots P(X_{n+1} = a_{n+1}|X_n = a_n) =$$
$$P_{a_1}P_{a_{1a_2}} \dots P_{a_{na_{n+1}}}.$$

فرض می‌کنیم برای $i, j = 1,2, \dots, c$ Ω_{ij} برابر تعداد دفعاتی باشد که فرایند از وضعیت i به وضعیت j رفته است و فرض می‌کنیم $\Omega = \{\Omega_{ij}\}$ یک ماتریس $c \times c$ باشد که آن را ماتریس تغییر وضعیت شمارشی می‌نامیم یعنی درایه‌های این ماتریس تعداد دفعاتی است که فرایند از یک وضعیت به وضعیت دیگر رفته است. از آنجایی که

$$P_{a_1}P_{a_{1a_2}} \dots P_{a_{na_{n+1}}} = P_{a_1} \prod_{ij} p_{\mathrm{ij}}{}^{\Omega_{ij}} . \qquad (1-1)$$

طرف راست تساوی $(1-1)$ مشابه با احتمال در توزیع چندجمله‌ای است، و احتمال رخ دادن هر مجموعه‌ای خاص از وضعیتها با شروع از حالت a_1 و داشتن ماتریس تغییر وضعیت شمارشی Ω به وسیله $(1-1)$ داده می‌شود [8]. لذا ضروری است که تعداد رخ دادهای این مجموعه را بشماریم تا بتوانیم توزیع آماره‌ی بسنده را پیدا کنیم. اگر $\Omega_{i.} = \sum_j \Omega_{ij}$، $\{\Omega_{i.}\}$ فراوانی شمارشهای a_1، ...، a_n است، و همچنین $n = \sum_{ij} \Omega_{ij} = \sum_i \Omega_{i.}$. می‌توانیم ببینیم که این ارتباط Ω و وضعیت آغازی به طور کامل وضعیت انتهایی را نشان می‌دهند. به عبارت دیگر با Ω و وضعیت آغازی روند فرایند معلوم می‌شود.

1-7-1. **قضیه**: فرض می‌کنیم Ω یک ماتریس $c \times c$ از اعداد صحیح مثبت که $n = \sum_{ij} \Omega_{ij}$ و $\Omega_{i.} - \Omega_{.j} = \delta_{iu} - \delta_{iv}$ ، $i = 1,2, \dots, c$ و برای برخی جفتهای u و v موجود باشند.

---

۱. P.Whittle

اگر $N_{\mathrm{u,v}}^{(\mathrm{n})}(\Omega)$ تعداد دنباله‌های از $a_{n+1},\dots,a_1$ با ماتریس تغییر وضعیت شمارشی Ω و $a_1 = u$ و $a_{n+1} = v$ باشد در این صورت

$$N_{\mathrm{u,v}}^{(\mathrm{n})}(\Omega) = \frac{\prod_{\mathrm{i}} \Omega_{i.}!}{\prod_{\mathrm{ij}} \Omega_{ij}!} \Omega_{vu}^{*} \ . \qquad (2-1)$$

که در آن Ω_{vu}^{*} (u,v) امین درایه از ماتریس $\Omega^{*} = \{\Omega_{ij}{}^{*}\}$ با درایه‌های زیر است.

$$\Omega_{ij}{}^{*} = \begin{cases} \delta_{ij} - {}^{\Omega_{ij}}\!/_{\Omega_{i.}} & \Omega_{i.} > o \\ \delta_{ij} & \Omega_{i.} = o \end{cases}$$

برهان: [8].

با توجه به $(1-1)$ و $(2-1)$ احتمال اینکه دنباله‌ی $\{x_1,\dots,x_{n+1}\}$ با ماتریس تغییر وضعیت شمارشی Ω و $x_1 = u$ و $x_{n+1} = v$ رخ دهد برابراست با :

$$P_u \Omega_{vu}^{*} \frac{\prod_{\mathrm{i}} \Omega_{i.}!}{\prod_{\mathrm{ij}} \Omega_{ij}!} \prod_{i,j} p_{\mathrm{ij}}{}^{\Omega_{ij}} \ . \qquad (3-1)$$

که عبارت فوق را فرمول ویتل می‌گوییم.

1-8. **روشهای خی دو**: روشهای خی دو روشی مناسب و کار آمد در حل مسئله تحلیل آماری زنجیرهای مارکوف که براساس تحققهای مارکوفی است و موارد شبیه چندجمله‌ای هستند می‌باشد. روشهای زیادی براساس خی دو ارائه شده‌اند برای مثال کرامر [21]. برای ساده کردن بحث ابتدا فرض می‌کنیم زنجیر مانا و ارگودیک باشد با نادیده گرفتن فاکتور $P_u \Omega_{vu}^{*}$ در فرمول ویتل (1-3) به طورکلی می‌توانیم ببینیم که احتمال از ماتریس تغییر وضعیت شمارشی Ω به صورت زیر است.

$$\prod_i \frac{\Omega_{i.}!}{\prod_j \Omega_{ij}!} \prod_j p_{ij}{}^{\Omega_{ij}} \ . \qquad (4-1)$$

اگر $\Omega_{i.}$ را به صورت Ω_i تعریف کنیم فرمول $(4-1)$ همانند احتمال بدست آوردن c فراوانی شمارشی $(\Omega_{i1},\dots,\Omega_{ic})$ در c نمونه مستقل از اندازه Ω_i از جامعه‌های چندجمله‌ای با احتمالهای خانه‌ای $(\mathrm{p}_{i1},\dots,\mathrm{p}_{ic})$ است. فرض می‌کنیم

$$\xi_{ij} = {(\Omega_{ij} - \Omega_i P_{ij})}\Big/{\sqrt{\Omega_i}} \,.$$

اگر این چندجمله‌ای معلوم باشد آنگاه c بردار تصادفی $\xi_i = (\xi_{i1}, \dots, \xi_{ic})$ از یک دیگر مستقل خواهند بود و کوواریانس ساختار ξ_i به صورت زیر خواهند بود

$$E(\xi_{ij}\xi_{il}) = \delta_{jl}\mathrm{p}_{ij} - \mathrm{p}_{ij}\mathrm{p}_{il} \,.$$

اگر Ω_i ها بزرگ باشند ξ_i دارای توزیع نرمال تقریبی خواهند بود، وقتی که n بزرگ می‌شود نتایج زیر را داریم[8].

1-8-1. **قضیه**: در حالتی که زنجیر مانا و ارگودیک است، بردار تصادفیc^2 بعدی، $\xi = (\xi_{ij})$ در توزیع به توزیع نرمال با ماتریس کوواریانس $(\lambda_{ij,jl})$ همگرا ست، که در آن

$$\lambda_{ij,jl} = \delta_{jk}(\delta_{jl}\mathrm{p}_{ij} - \mathrm{p}_{ij}\mathrm{p}_{il}) \,.$$

برهان: [8].

با فرض وجود قضیه‌ی (1-8-1) ، از نظریه خی دو معمولی در هر آماره داریم

$$\sum_j \frac{(\Omega_{ij} - \Omega_i p_{ij})^2}{\Omega_i p_{ij}} \qquad i = 1,2, \dots, \mathrm{c} \,. \qquad (5-1)$$

به طور مجانبی دارای توزیع خی دو در $(5-1)$ است و درایه j محدود و$\mathrm{p}_{ij} > 0$ است. اگر تعداد اینها d_i باشد آنگاه توزیع خی دو دارای $d_\mathrm{i} - 1$ درجه آزادی است به علاوه c آماره به طور مجانبی مستقل هستند، و مجموع آنها به صورت زیر است.

$$\sum_{ij} \frac{(\Omega_{ij} - \Omega_i \mathrm{p}_{ij})^2}{\Omega_i \mathrm{p}_{ij}} \qquad (6-1)$$

به طور مجانبی دارای توزیع خی دو با $d - c$ درجه آزادی است که $d = \sum_{i=1}^{c} d_\mathrm{i}$ ($\sum_{i=1}^{c} d_\mathrm{i}$ پارامتر داریم که c تای آن را برآورد می‌کنیم لذا درجه آزادی $d - c$ می‌شود.) . آماره‌ی $(6-1)$ اولین بار توسط بارتلت [6] در نظر گرفته شد که یک معیار برای برازش نیکویی از نمونه با استفاده از