

استنباط آماری زنجیرهای مارکوف

* غلامحسین شاهکار

نظر بگیرید. (فرض کنید z_i تعداد تغییر وضعیت‌های یک مرحله‌ای مشاهده شده از وضعیت z به z_i و N تعداد کل مشاهدات باشد). گیریم

$$\sum_{j=1}^m n_{ij} = n_{i \cdot}, \quad \sum_{i=1}^m n_{ij} = n_{j \cdot}, \\ \sum_{i=1}^m \sum_{j=1}^m n_{ij} = \sum_{i=1}^m n_{i \cdot} = n..$$

ویتل (۱۹۵۵) احتمال مشاهده تغییر وضعیتها را به صورت زیر به دست آورد:

$$f(n_{ij}) = \prod_{i=1}^m \left[g(n_{i \cdot}) \frac{(n_{i \cdot})!}{\prod_{j=1}^m (n_{ij})!} \prod_{j=1}^m p_{ij}^{n_{ij}} \right] \quad (1)$$

می‌توان لگاریتم تابع درستنمایی را به صورت زیر نوشت

$$L(p_{ij}) = c + \sum_{i=1}^m \sum_{j=1}^m n_{ij} \log(p_{ij}) \quad (2)$$

که در آن c شامل تمام جملات مستقل از p_{ij} هاست. چون $1 = \sum_{j=1}^m p_{ij}$ لذا می‌توان (۲) را چنین نوشت:

$$L(p_{ij}) = c + \sum_{i=1}^m \sum_{j=1}^{m-1} n_{ij} \log(p_{ij}) + \sum_{i=1}^m n_{im} \log \left[1 - \sum_{j=1}^{m-1} p_{ij} \right] \quad (3)$$

فرض کنید τ مقدار مشخص شده‌ای برای z باشد. برای هر مقدار $k = 1, 2, \dots, m-1$, برآوردهای درستنمایی ماکسیمم $\hat{p}_{\tau k}$ از حل معادلات زیر به دست می‌آید

چکیده

مسائل استنباط آماری زنجیرهای مارکوف مانند برآورد و آزمون فرضها در زمینه‌های مختلف کاربردی اهمیت ویژه دارند و متأسفانه کمتر مورد توجه مؤلفان قرار گرفته‌اند. در این مقاله ضمن شرح روش درستنمایی ماکسیمم برای برآورد احتمالهای تغییر وضعیت زنجیرهای مارکوف، در نحوه آزمون رتبه یک زنجیر مارکوف نیز بحث می‌کیم.

۱ مقدمه

مسائل استنباط آماری زنجیرهای مارکوف مانند برآورد و آزمون فرضها را محققانی نظری بارتلت^۱، ویتل^۲، اندرسون^۳، گودمن^۴ و بیلینگزل^۵ نه تنها به دلیل جالب بودن جنبه‌های نظری آنها بلکه به دلیل کاربردشان در زمینه‌های مختلف بررسی کرده‌اند. برای برآورد احتمالهای تغییر وضعیت زنجیرهای مارکوف مرتبه اول و بالاتر، روش‌هایی که جدیدترین آنها مبتنی بر شیوه‌های برنامه‌ریزی خطی و درجه دوم برای حصول برآوردهای کمترین مرباعات است ارائه شده‌اند (لی^۶ و همکاران^۷ (۱۹۷۰) را ببینید). در اینجا روش درستنمایی ماکسیمم را برای برآورد احتمالهای تغییر وضعیت با استفاده از مشاهدات به عمل آمده معرفی می‌کنیم. و براساس این برآورد چندین آزمون را نیز ارائه می‌دهیم.

۲ برآورده درستنمایی ماکسیمم

یک زنجیر مارکوف همگن با تعدادی متناهی از وضعیتها متلاً m وضعیت و با ماتریس احتمال تغییر وضعیت $(z_i p_{ij})_{i,j=1,2,\dots,m}$ را در

*) دکتر غلامحسین شاهکار، گروه آمار دانشگاه فردوسی مشهد

1) Bartlett 2) Whittle 3) Anderson 4) Goodman 5) Billingsley 6) Lee

$$\text{آماره} \\ \sum_{j=1}^m \frac{n_{ij}(\hat{p}_{ij} - p_{ij})^2}{p_{ij}}, \quad i = 1, 2, \dots, m \quad (11)$$

دارای توزیع مجانبی χ^2 با $m-1$ درجه آزادی است. البته باید زو_iهایی را که برابر صفرند خارج کرده و از درجه آزادی به اندازه تعداد آنها کم کنیم. با جمعبندی آماره بالا روی تمام مقادیر p_{ij} می‌توان آزمونی برای تمام زو_iهایی دست آورد. در این حالت آماره

$$\sum_{i=1}^m \sum_{j=1}^m \frac{n_{ij}(\hat{p}_{ij} - p_{ij})^2}{p_{ij}} \quad (12)$$

دارای توزیع مجانبی χ^2 با $(m-1)(m-2)$ درجه آزادی است (در اینجا نیز لازم است از تعداد درجات آزادی به اندازه تعداد زو_iهایی که به ازای $i, j = 1, 2, \dots, m$ برابر صفر هستند کم کنیم).

با استفاده از نسبت درستنمایی، آماره مناسب آزمون برای فرض H_0 عبارت است از

$$\lambda = \prod_{i=1}^m \prod_{j=1}^m \left[\frac{\hat{p}_{ij}}{p_{ij}} \right]^{n_{ij}}$$

و تحت فرض صفر آماره

$$-2 \log \lambda = -2 \sum_{i=1}^m \sum_{j=1}^m n_{ij} \log \frac{n_{ij}}{n_{i\cdot} p_{ij}} \quad (13)$$

دارای توزیع مجانبی χ^2 با $(m-1)(m-2)$ درجه آزادی است. قبل از معرفی آزمون دوم اندرسن، مرتبه یک فرایند تصادفی را تعریف می‌کنیم. فرایندی را از مرتبه k می‌گوییم هرگاه در توزیع شرطی X_n به شرط معلوم بودن X_0, X_1, \dots, X_{n-1} فقط k مقدار آخر در این توزیع شرطی مؤثر باشد و بتوان از متغیرهای قبل از آن صرف نظر کرد. به زیان نمادها خواهیم داشت

$$P(X_n = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) =$$

$$P(X_n = y | X_{n-1} = x_{n-1}, \dots, X_{n-k} = x_{n-k})$$

برای آزمون مرتبه یک زنجیر مارکوف نیز می‌توان از برآوردهای درستنمایی ماکسیمم استفاده کرد. آماره آزمون برای آزمون این فرض که زنجیر از مرتبه صفر است، یعنی به ازای تمام مقادیر p_{ij} می‌توان آن را با $H_0 : p_{ij} = p_{ij}^*$ در برابر این فرض که زنجیر از مرتبه یک است عبارت است از

$$\lambda = \prod_{i=1}^m \prod_{j=1}^m \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}^*} \right]^{n_{ij}} \quad (14)$$

که در آن

$$\hat{p}_j = \frac{n_{\cdot j}}{n_{\cdot \cdot}}, \quad \hat{p}_{ij} = \frac{n_{ij}}{n_{i\cdot}}$$

$$\frac{\partial L(p_{rk})}{\partial p_{rk}} = 0, \quad k = 1, 2, \dots, m-1 \quad (4)$$

از (۳) و (۴) نتیجه می‌شود

$$\frac{n_{rk}}{p_{rk}} - \frac{n_{rm}}{1 - \sum_{j=1}^{m-1} p_{rj}} = 0, \quad k = 1, 2, \dots, m-1 \quad (5)$$

حال مقدار مشخصی برای k مثلاً s را در نظر بگیرید، داریم

$$\begin{aligned} \frac{n_{rs}}{p_{rs}} &= \frac{n_{rk}}{p_{rk}} \\ &= \frac{n_{rm}}{1 - \sum_{j=1}^{m-1} p_{rj}}, \quad k = 1, 2, \dots, s, \dots, m-1 \end{aligned} \quad (6)$$

بنابراین

$$1 - \sum_{j=1}^{m-1} p_{rj} = \frac{n_{rm}}{n_{rs}} p_{rs} \quad (7)$$

$$p_{rk} = \frac{n_{rk}}{n_{rs}} p_{rs}, \quad k = 1, 2, \dots, s, \dots, m-1 \quad (8)$$

اگر طرفین (۸-۲) را روی مقادیر $1, 2, \dots, m-1$ جمع کرده و نتیجه را به (۷-۲) اضافه کنیم به دست می‌آوریم

$$1 = \frac{\sum_{k=1}^m n_{rk}}{n_{rs}} p_{rs}$$

و بنابراین برآورد \hat{p}_{rs} را همان طور که به طور شهودی نیز انتظار داریم به صورت زیر نتیجه می‌گیریم

$$\hat{p}_{rs} = \frac{n_{rs}}{\sum_{k=1}^m n_{rk}} \quad (9)$$

اما r و s دو مقدار دلخواه برای \hat{p}_{rs} هستند. پس برای هر جفت دلخواه از مقادیر $1, 2, \dots, m-1$ برابر \hat{p}_{rs} درستنمایی ماکسیمم p_{rs} عبارت است از

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{k=1}^m n_{ik}} = \frac{n_{ij}}{n_{i\cdot}} \quad (10)$$

۳ آزمون فرضها

دو آزمون زیر به وسیله اندرسن برای برآوردهای بالا ارائه شد:

۱) می‌خواهیم این فرض را که مشاهدات، مقادیر محقق یک زنجیر مارکوف با ماتریس احتمال تغییر وضعیت P^* هستند آزمون کنیم. به عبارت دیگر فرض صفر عبارت است از

$$H_0 : P = P^*$$

$-2 \log \lambda_{r-1,r} = 2 \sum_i \dots \sum_l n_{ij\dots kl} \log \frac{(n_{ij\dots kl})(n_{\dots\dots\dots})}{(n_{(n_{ij\dots k})})(n_{j\dots kl})}$

که متغیری تصادفی با توزیع χ^2 و با $(m-1)^2$ درجه آزادی است. توجه کنید که درستی فرض H_k برای $r < k$ درستی فرض H_r را نتیجه می‌دهد. $\lambda_{k,r}$ را نسبت درستنمایی ماکسیمم برای آزمون فرض در مقابل فرض H_r (یعنی فرض اینکه زنجیر از مرتبه k است در برای این فرض که زنجیر از مرتبه r است) می‌گیریم. پس به دست می‌آوریم

$$\lambda_{k,r} = \prod_{j=k+1}^r \lambda_{k,j} \quad (17)$$

و بنابراین

$$-2 \log \lambda_{k,r} = -2 \log \lambda_{k,k+1} - \dots - 2 \log \lambda_{r-1,r}, \quad k < r \quad (18)$$

فرض کنید متغیرهای $-2 \log \lambda_{r-1,r}, -2 \log \lambda_{r-2,r}, \dots, -2 \log \lambda_{1,r}$ به طور مجانبی مستقل باشند، پس با قبول فرض H_k نتیجه می‌شود که $-2 \log \lambda_{k,r}$ دارای توزیع χ^2 با تعداد درجات آزادی زیر است

$$\begin{cases} \nabla m^{r+1} - \nabla m^{k+1}, & k \geq 0 \\ \nabla m^{r+1}, & k = -1 \end{cases}$$

که در آن

$$\nabla m^r = m^r - m^{r-1} \quad r \geq 1$$

در استفاده از این شیوه به عنوان یک روش تصمیم‌گیری، چگونگی انتخابتابع مناسب زیان مطرح می‌شود. توابع زیانی که آنها را در نظریه کلاسیک آزمون فرضها در نظر می‌گیرند احتمال قبول فرض نادرست یا احتمال رد فرض درست تعريف می‌کنند.

ثانگ بر اساس روش میزان اطلاع آکائیک تابع زیان را به صورت زیر پیشنهاد می‌کند

$$R(k) = -2 \log \lambda_{k,m} - 2(\nabla S^{m+1} - \nabla S^{k+1})$$

که در آن m بالاترین مرتبه در نظر گرفته شده برای مدل و k مرتبه برآزنده شده است. تقریب خوبی برای مرتبه یک زنجیر مارکوف براساس برآورد مینیمم اطلاع آکائیک مقداری از k است که $R(k)$ را نسبت به تمام مرتبه‌های در نظر گرفته شده مینیمم کند. توجه کنید که $R(m) = 0$.

گابریل⁷⁾ و نیومن⁸⁾ وضعیت بارندگی و عدم بارندگی هوا را به صورت یک زنجیر مارکوف دو وضعیتی توصیف کرده‌اند. آنها هواهای غیر بارانی را با صفر و هوای بارانی را با یک شان داده‌اند. مثال زیر مدل مشابهی برای داده‌های هواشناسی است که مدهنی⁹⁾ (۱۹۷۶) از آن داده است.

تحت فرض صفر، آماره

$$-2 \log \lambda = 2 \sum_i \sum_j n_{ij} \log \frac{n_{ij}}{n_{j\dots i}}$$

دارای توزیع مجانبی χ^2 با $(m-1)^2$ درجه آزادی است. به همین نحو می‌توان برای آزمون این فرض که زنجیر از مرتبه اول است در برای این فرض که زنجیر از مرتبه دوم است آزمونی ارائه کرد.

البته این آزمونها ثابت نمی‌کنند که زنجیر مشاهده شده از مرتبه اول است، اما منطقی را برای استفاده از مدل مارکوف به دست می‌دهند. بر پایه برآوردهای درستنمایی ماکسیمم آزمونهای مشابهی با استفاده از نسبت درستنمایی برای (الف) آزمون مانایی (یا همگن زمانی) یک زنجیر مارکوف پنهانی آزمون $p_{ij}(t) = p_{ij}$ و همچنین (ب) برای آزمون اینکه آیا چندین نمونه از زنجیر مارکوفی با مرتبه معین آمدند ارائه شده است:

۴ تعیین مرتبه یک زنجیر مارکوف به وسیله برآورد مینیمم میزان اطلاع آکائیک

تونگ¹⁰⁾ (۱۹۷۵) شیوه‌ای برای تعیین مرتبه یک زنجیر مارکوف به کمک میزان اطلاع آکائیک¹¹⁾ ارائه کرده است. میزان اطلاع آکائیک به صورت زیر تعریف می‌شود.

= میزان اطلاع آکائیک

$$(15) \quad \text{تعداد پارامترهای مدل} + (\text{درستنمایی ماکسیمم}) - 2 \log$$

این آماره اندازه‌ای برای انحراف مدل برآزنده شده از ساختار واقعی را به دست می‌دهد. با داشتن چندین مدل، این شیوه به قبول مدلی منجر می‌شود که میزان اطلاع آکائیک برای آن مینیمم باشد و بنابراین آن را برآورد مینیمم میزان اطلاع آکائیک می‌گویند. (این شیوه کوششی است جهت ایجاد موازنه میان برآش بالاکه نیاز به پارامترهای بیشتر دارد و برآش پایین که موجب افزایش واریانس برای مانده می‌شود).

احتمالهای تغییر وضعیت زنجیر مرتبه r را با $p_{ij\dots k\dots l}, i, j = 1, 2, \dots, m, p_{ij\dots k\dots l}$ نشان می‌دهیم که در آن m را کل تعداد وضعیت‌های زنجیر فرض کرده‌ایم.

در اینجا برآورد درستنمایی ماکسیمم عبارت است از

$$(16) \quad p_{ij\dots kl} = \frac{n_{ij\dots kl}}{\sum_l p_{ij\dots kl}}$$

که در آن $n_{ij\dots k\dots l} = \sum_l p_{ij\dots kl}$. فرض آزمون عبارت است از

$$H_{r-1} : p_{ij\dots kl} = p_{j\dots kl}, \quad i = 1, 2, \dots, m$$

(یعنی زنجیر از مرتبه $1 \leq r$ است در برای این فرض که زنجیر از مرتبه r است). آماره آزمون عبارت است از

و می‌توان توزیع حدی را به صورت زیر به دست آورد

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = (0,60377, 0,39623)$$

می‌خواهیم فرض

$$H_0 : p_{ij} = p_j, \quad i, j = 1, 2$$

یعنی این فرض را که زنجیر از مرتبه صفر است در برابر این فرض که زنجیر از مرتبه یک است آزمون کنیم. داریم

$$\begin{aligned} -2 \log \lambda &= 2 \sum_{i=1}^2 \sum_{j=1}^2 n_{ij} \log \frac{N n_{ij}}{n_i \cdot n_j} \\ &= 2 \left\{ 175 \log \frac{367 \times 175}{223 \times 224} + 49 \log \frac{368 \times 49}{140 \times 224} \right. \\ &\quad \left. + 48 \log \frac{368 \times 48}{223 \times 144} + 96 \log \frac{368 \times 96}{140 \times 144} \right\} \\ &= 53,28 \end{aligned}$$

$P(\chi^2 \geq 53,28)$ با یک درجه آزادی سیار کوچک است و بنابراین فرض

صفر از مرتبه صفر بودن زنجیر را رد می‌کنیم. همچنین مایلیم فرض

$$H_0 : \mathbf{P} = \mathbf{P}^* = \begin{pmatrix} \frac{7}{15} & \frac{8}{15} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

را آزمون کنیم که در آن \mathbf{P}^* ماتریس احتمال تغییر وضعیت برای روزهای غیر بارانی و بارانی طی ماههای زوئن تا اوتو است. داریم

$$\sum_{i=1}^2 \sum_{j=1}^2 \frac{n_{ij}(p_{ij} - p_{ij}^*)}{p_{ij}^*} = 105,07$$

$P(\chi^2 \geq 105,07)$ با دو درجه آزادی خیلی کوچک است، و بنابراین فرض H_0 را نیز رد می‌کنیم.

مثال داده‌های زیر مشاهدات به عمل آمده از وضعیت بارندگی و عدم بارندگی روزهای مختلف در ناحیه گاهاتی هندوستان است. وضعیت غیر بارانی را با صفر و وضعیت بارانی را با یک نشان داده‌ایم. مدت زمان بررسی از آخرین روز ماه فوریه (یعنی قبل از آغاز موسیم بارندگی) تا آخرین روز ماه مه طی چهار سال متولی از ۱۹۶۷ تا ۱۹۷۰ بوده است و N کل تعداد تغییر وضعیتها طی چهار سال، یا $368 = 4 \times 92$ روز است. تغییر وضعیتهای گزارش شده عبارت اند از

	۰	۱	جمع
۰	۱۷۵	۴۹	۲۲۴
۱	۴۸	۹۶	۱۴۴
جمع	۲۲۳	۱۴۵	۳۶۸

پس به موجب (۱۰) ماتریس تغییر وضعیت $(p_{ij}) = \mathbf{P}$ را به صورت زیر برآورد می‌کنیم

$$\mathbf{P} = \begin{pmatrix} \frac{175}{224} & \frac{49}{224} \\ \frac{48}{144} & \frac{96}{144} \end{pmatrix}$$

داریم

$$\mathbf{P}^0 = \begin{pmatrix} 0,61091 & 0,38909 \\ 0,59289 & 0,40711 \end{pmatrix}$$

$$\mathbf{P}^{10} = \begin{pmatrix} 0,30390 & 0,39610 \\ 0,60358 & 0,39642 \end{pmatrix}$$

$$\mathbf{P}^{100} = \begin{pmatrix} 0,60377 & 0,39623 \\ 0,60377 & 0,39623 \end{pmatrix}$$

مراجع

- [1] T. W. Anderson and L. A. Goodman. "Statistical Inference about Markov chains," Ann. Math. Stat. 28, 89-110 (1957).
- [2] P. Billingsley, "Statistical methods in Markov chains". Ann. Math. Stat., 32, 12-40 (1961).
- [3] P. Billingsley, "Statistical Inference for Markov Processes", University of Chicago Press. Chicago (1961).
- [4] T. C. Lee, G. C. Judge and A. Zellner, Estimating the Parameters of the Markov Probability Model for Aggregate time series Data, North Holland, Amsterdam, (1970).
- [5] J. Medhi, "A Markov chain model for the Occurrences of dry and wet days", Ind. J. Met. Hydrol. Geophys., 27, 431-435 (1976).
- [6] H. Tong, "Determination of the order of a Markov chain by Akaike's information criterion," J. Appl. Prob., 21, 488-497 (1975).
- [7] P. Whittle, "Some distributions and moment formulae for The Markov chains," J.R.S.S.B 17, 235-242 (1955).