استنباط آماری زنجیرهای مارکوف

غلامحسین شاهکار *

چکیده

مسائل استنباط آماری زنجیرهای مارکوف مانند برآورد و آزمون فرضها در زمینه‌های مختلف کاربردی اهمیّت 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 نحوهٔ آزمون رتبهٔ یک زنجیر مارکوف نیز بحث می‌کنیم.

۱ مقدمه

مسائل استنباط آماری زنجیرهای مارکوف مانند برآورد و آزمون فرضها را محققانی نظیر بارتلت[1]، ویتل[2]، اندرسن[3]، گودمن[4] و بیلینگزلی[5] 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 مربعات است ارائه شده‌اند (لی[6] و همکاران (۱۹۷۰) را ببینید). در اینجا روش درستنمایی ماکسیمم را برای برآورد احتمالهای تغییر وضعیت با استفاده از مشاهدات به‌عمل آمده معرفی می‌کنیم. و بر اساس این برآورد چندین آزمون را نیز ارائه می‌دهیم.

۲ برآورد درستنمایی ماکسیمم

یک زنجیر مارکوف همگن با تعدادی متناهی از وضعیتها مثلاً m وضعیت و با ماتریس احتمال تغییر وضعیت $\mathbf{P}=(p_{ij})$، $i,j=1,2,\dots,m$ را در نظر بگیرید. (فرض کنید n_{ij} تعداد تغییر وضعیتهای یک مرحله‌ای مشاهده شده از وضعیت i به j، و $N+1$ تعداد کل مشاهدات باشد.) گیریم

$$\sum_{j=1}^{m} n_{ij} = n_{i.}, \qquad \sum_{i=1}^{m} n_{ij} = n_{.j},$$
$$\sum_{i=1}^{m}\sum_{j=1}^{m} n_{ij} = \sum_{i=1}^{m} n_{i.} = n_{..}$$

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

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

می‌توان لگاریتم تابع درستنمایی را به صورت زیر نوشت

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

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

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

فرض کنید r مقدار مشخص شده‌ای برای i باشد. برای هر مقدار $k = 1, 2, \dots, m-1$، برآورد درستنمایی ماکسیمم $\hat{p}_{rk}$ از حل معادلات زیر به دست می‌آید

*) دکتر غلامحسین شاهکار، گروه آمار دانشگاه فردوسی مشهد

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

$$\frac{\partial L(p_{rk})}{\partial p_{rk}} = ۰, \qquad k = ۱,۲,\dots,m-۱ \qquad (۴)$$

از (۳) و (۴) نتیجه می‌شود

$$\frac{n_{rk}}{p_{rk}} - \frac{n_{rm}}{۱ - \sum_{j=۱}^{m-۱} p_{rj}} = ۰, \qquad k = ۱,۲,\dots,m-۱ \qquad (۵)$$

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

$$\frac{n_{rs}}{p_{rs}} = \frac{n_{rk}}{p_{rk}} = \frac{n_{rm}}{۱ - \sum_{j=۱}^{m-۱} p_{rj}}, \qquad k = ۱,۲,\dots,s,\dots,m-۱ \qquad (۶)$$

بنابراین

$$۱ - \sum_{j=۱}^{m-۱} p_{rj} = \frac{n_{rm}}{n_{rs}} p_{rs} \qquad (۷)$$

و

$$p_{rk} = \frac{n_{rk}}{n_{rs}} p_{rs}, \qquad k = ۱,۲,\dots,s,\dots,m-۱ \qquad (۸)$$

اگر طرفین (۲-۸) را روی مقادیر $k = ۱,۲,\dots,m-۱$ جمع کرده و نتیجه را به (۲-۷) اضافه کنیم به دست می‌آوریم

$$۱ = \frac{\sum_{k=۱}^{m} n_{rk}}{n_{rs}} p_{rs}$$

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

$$\hat{p}_{rs} = \frac{n_{rs}}{\sum_{k=۱}^{m} n_{rk}} \qquad (۹)$$

اما r و s دو مقدار دلخواه برای i و j هستند. پس برای هر جفت دلخواه از مقادیر $i,j = ۱,۲,\dots,m-۱$ برآورد درستنمایی ماکسیمم p_{ij} عبارت است از

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{k=۱}^{m} n_{ik}} = \frac{n_{ij}}{n_{i.}} \qquad (۱۰)$$

۳ آزمون فرضها

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

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

$$H_{\circ} : \mathbf{P} = \mathbf{P}^{\circ}$$

آمارهٔ

$$\sum_{j=۱}^{m} \frac{n_{i.}(\hat{p}_{ij} - p_{ij}^{\circ})^{۲}}{p_{ij}^{\circ}}, \qquad i = ۱,۲,\dots,m \qquad (۱۱)$$

دارای توزیع مجانبی $\mathcal{X}^{۲}$ با $m-۱$ درجهٔ آزادی است. البته باید p_{ij}°هایی را که برابر صفرند خارج کرده و از درجهٔ آزادی به اندازه تعداد آنها کم کنیم. با جمعبندی آمارهٔ بالا روی تمام مقادیر i می‌توان آزمونی برای تمام $\hat{p}_{ij}$ها به دست آورد. در این حالت آمارهٔ

$$\sum_{i=۱}^{m}\sum_{j=۱}^{m} \frac{n_{i.}(\hat{p}_{ij} - p_{ij}^{\circ})^{۲}}{p_{ij}^{\circ}} \qquad (۱۲)$$

دارای توزیع مجانبی $\mathcal{X}^{۲}$ با $m(m-۱)$ درجهٔ آزادی است (در اینجا نیز لازم است از تعداد درجات آزادی به اندازهٔ تعداد p_{ij}°هایی که به ازای $i,j = ۱,۲,\dots,m$ برابر صفر هستند کم کنیم).

با استفاده از نسبت درستنمایی، آمارهٔ مناسب آزمون برای فرض $H_{\circ}$ عبارت است از

$$\lambda = \prod_{i=۱}^{m}\prod_{j=۱}^{m}\left[\frac{\hat{p}_{ij}}{p_{ij}^{\circ}}\right]^{n_{ij}}$$

و تحت فرض صفر آمارهٔ

$$-۲\log\lambda = -۲\sum_{i=۱}^{m}\sum_{j=۱}^{m} n_{ij}\log\frac{n_{ij}}{n_{i.}p_{ij}^{\circ}} \qquad (۱۳)$$

دارای توزیع مجانبی $\mathcal{X}^{۲}$ با $m(m-۱)$ درجهٔ آزادی است.

قبل از معرفی آزمون دوم اندرسن، مرتبهٔ یک فرایند تصادفی را تعریف می‌کنیم. فرایندی را از مرتبهٔ k می‌گوییم هرگاه در توزیع شرطی X_n به شرط معلوم بودن $X_{\circ},\dots,X_{n-۱}$ فقط k مقدار آخر در این توزیع شرطی مؤثر باشند و بتوان از متغیرهای قبل از آن صرف‌نظر کرد. به زبان نمادها خواهیم داشت

$$P(X_n = y | X_{\circ} = x_{\circ},\dots,X_{n-۱} = x_{n-۱}) = P(X_n = y | X_{n-۱} = x_{n-۱},\dots,X_{n-k} = x_{n-k})$$

۲) برای آزمون مرتبهٔ یک زنجیر مارکوف نیز می‌توان از برآوردهای درستنمایی ماکسیمم استفاده کرد. آمارهٔ آزمون برای آزمون این فرض که زنجیر از مرتبهٔ صفر است، یعنی به ازای تمام مقادیر i، $H_{\circ} : p_{ij} = p_j$ در برابر این فرض که زنجیر از مرتبهٔ یک است عبارت است از

$$\lambda = \prod_{i=۱}^{m}\prod_{j=۱}^{m}\left[\frac{\hat{p}_i}{\hat{p}_{ij}}\right]^{n_{ij}} \qquad (۱۴)$$

که در آن

$$\hat{p}_j = \frac{n_{.j}}{n_{..}}, \qquad \hat{p}_{ij} = \frac{n_{ij}}{n_{i.}}$$

تحت فرض صفر، آماره

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

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 برابر این فرض که زنجیر از مرتبهٔ دوم است آزمونی ارائه کرد.

البته این آزمونها ثابت نمی‌کنند که زنجیر مشاهده شده از مرتبهٔ اول است، اما منطقی را برای استفاده از مدل مارکوف به دست می‌دهند. بر پایهٔ برآوردهای درستنمایی ماکسیمم آزمونهای مشابهی با استفاده از نسبت درستنمایی برای (الف) آزمون مانایی (یا همگن زمانی) یک زنجیر مارکوف یعنی آزمون $H_0 : p_{ij}(t) = p_{ij}$ و همچنین (ب) برای آزمون اینکه آیا چندین نمونه از زنجیر مارکوفی با مرتبه معین آمده‌اند ارائه شده است.

۴ تعیین مرتبهٔ یک زنجیر مارکوف به وسیلهٔ برآورد مینیمم میزان اطلاع آکائیک

تونگ[7] (۱۹۷۵) شیوه‌ای برای تعیین مرتبهٔ یک زنجیر مارکوف به کمک میزان اطلاع آکائیک[8] ارائه کرده است. میزان اطلاع آکائیک به صورت زیر تعریف می‌شود.

= میزان اطلاع آکائیک

(۱۵) (تعداد پارامترهای مدل)۲ + (درستنمایی ماکسیمم)$2\log$ − .

این آماره اندازه‌ای برای انحراف مدل برازانده 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 برازش بالا که نیاز به پارامترهای بیشتر دارد و برازش پایین که موجب افزایش واریانس برای مانده می‌شود).

احتمالهای تغییر وضعیت زنجیر مرتبهٔ r را با $p_{ij\ldots kl}$، $i,j = 1,2,\ldots,m$ نشان می‌دهیم که در آن m را کل تعداد وضعیتهای زنجیر فرض کرده‌ایم. در اینجا برآورد درستنمایی ماکسیمم عبارت است از

$$\hat{p}_{ij\ldots kl} = \frac{n_{ij\ldots kl}}{n_{ij\ldots k.}} \qquad (۱۶)$$

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

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

(یعنی زنجیر از مرتبهٔ $r-1$ است در برابر این فرض که زنجیر از مرتبه r است.) آمارهٔ آزمون عبارت است از

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

که متغیری تصادفی با توزیع χ^2 و با $m^{r-1}(m-1)^2$ درجه آزادی است.

توجه کنید که درستی فرض H_k برای $k < r$ درستی فرض H_r را نتیجه می‌دهد. $\lambda_{k,r}$ را نسبت درستنمایی ماکسیمم برای آزمون فرض H_k در مقابل فرض H_r (یعنی فرض اینکه زنجیر از مرتبهٔ k است در برابر این فرض که زنجیر از مرتبه r است) می‌گیریم. پس به دست می‌آوریم

$$\lambda_{k,r} = \prod_{j=k+1}^{r}\lambda_{k,j} \qquad (۱۷)$$

و بنابراین

$$-2\log\lambda_{k,r} = -2\log\lambda_{k,k+1} - \cdots - 2\log\lambda_{r-1,r},\ k<r \qquad (۱۸)$$

فرض کنید متغیرهای $-2\log\lambda_{r-1,r}$ برای $r = 0,1,2,\ldots$ 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

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

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

تانگ 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 مرتبهٔ در نظرگرفته شده برای مدل و k مرتبهٔ برازانده 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 $R(m) = 0$.

گابریل[9] و نیومن[10] وضعیت بارندگی و عدم بارندگی هوا را به صورت یک زنجیر مارکوف دو وضعیتی توصیف کرده‌اند. آنها هوای غیر بارانی را با صفر و هوای بارانی را با یک نشان داده‌اند. مثال زیر مدل مشابهی برای داده‌های هواشناسی است که مدهی[11] (۱۹۷۶) ارائه داده است.

---

7) Tong 8) Akaike 9) Gabriel 10) Newman 11) Medhi

مثال داده‌های زیر مشاهدات به عمل آمده از وضعیت بارندگی و عدم بارندگی روزهای مختلف در ناحیه گاهاتای 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 موسم بارندگی) تا آخرین روز ماه مه طی چهار سال متوالی از ۱۹۶۷ تا ۱۹۷۰ بوده است و N کل تعداد تغییر وضعیتها طی چهار سال، یا $۳۶۸ = ۹۲ \times ۴$ روز است. تغییر وضعیتهای گزارش شده عبارت‌اند از

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

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

$$\begin{array}{c|cc} & ۰ & ۱ \\ \hline ۰ & \frac{۱۷۵}{۲۲۴} & \frac{۴۹}{۲۲۴} \\ ۱ & \frac{۴۸}{۱۴۴} & \frac{۹۶}{۱۴۴} \end{array}$$

داریم

$$\mathbf{P}^۵ = \begin{pmatrix} ۰/۶۱۰۹۱ & ۰/۳۸۹۰۹ \\ ۰/۵۹۲۸۹ & ۰/۴۰۷۱۱ \end{pmatrix}$$

$$\mathbf{P}^{۱۰} = \begin{pmatrix} ۰/۳۰۳۹۰ & ۰/۳۹۶۱۰ \\ ۰/۶۰۳۵۸ & ۰/۳۹۶۴۲ \end{pmatrix}$$

$$\mathbf{P}^{۲۰} = \begin{pmatrix} ۰/۶۰۳۷۷ & ۰/۳۹۶۲۳ \\ ۰/۶۰۳۷۷ & ۰/۳۹۶۲۳ \end{pmatrix}$$

و می‌توان توزیع حدی را به صورت زیر به دست آورد

$$\lim_{n\to\infty} \mathbf{P}^n = (۰/۶۰۳۷۷, ۰/۳۹۶۲۳)$$

می‌خواهیم فرض

$$H_۰ : p_{ij} = p_j, \qquad i,j = ۱,۲$$

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

$$\begin{aligned} -۲\log\lambda &= ۲\sum_{i=۱}^{۲}\sum_{j=۱}^{۲} n_{ij}\log\frac{Nn_{ij}}{n_{i.}n_{.j}} \\ &= ۲\left\{۱۷۵\log\frac{۳۶۷\times۱۷۵}{۲۲۳\times۲۲۴} + ۴۹\log\frac{۳۶۸\times۴۹}{۱۴۵\times۲۲۴}\right\} \\ &+ ۴۸\log\frac{۳۶۸\times۴۸}{۲۲۳\times۱۴۴} + ۹۶\log\frac{۳۶۸\times۹۶}{۱۴۵\times۱۴۴} \\ &= ۵۳/۲۸ \end{aligned}$$

$P(\mathcal{X}^۲ \geq ۵۳/۲۸)$ با یک درجه آزادی بسیار کوچک است و بنابراین فرض صفر از مرتبهٔ صفر بودن زنجیر را رد می‌کنیم. همچنین مایلیم فرض

$$H_۰ : \mathbf{P} = \mathbf{P}^۰ = \begin{pmatrix} \frac{۷}{۱۵} & \frac{۸}{۱۵} \\ \frac{۱}{۵} & \frac{۴}{۵} \end{pmatrix}$$

را آزمون کنیم که در آن $\mathbf{P}^۰$ ماتریس احتمال تغییر وضعیت برای روزهای غیر بارانی و بارانی طی ماههای ژوئن تا اوت است. داریم

$$\sum_{i=۱}^{۲}\sum_{j=۱}^{۲}\frac{n_{i.}(p_{ij}-p_{ij}^۰)}{p_{ij}^۰} = ۱۰۵/۰۷$$

$P(\mathcal{X}^۲ \geq ۱۰۵/۰۷)$ با دو درجه آزادی خیلی کوچک است، و بنابراین فرض $H_۰$ را نیز رد می‌کنیم.

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 Occurences 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).