بسمه تعالی

بازشناسی آماری الگو

دانشگاه صنعتی خواجه نصیرالدین طوسی — تمرین سری چهارم — تاریخ تحویل: ۱۴۰۴/۸/۱۷ — آزمایشگاه بینایی ماشینی و پردازش تصاویر پزشکی

۱- (اختیاری) سکه‌ای را ۱۰ بار پرتاب می‌کنیم. نتیجه آن T T T H T T T H T T است.

الف. می‌دانیم که نتیجه پرتاب سکه از توزیع برنولی پیروی می‌کند. پارامتر این توزیع (μ) را با استفاده از روش MLE تخمین بزنید.

ب. با توجه به داده‌های فوق مقدار عددی این پارامتر برای سکه فوق چقدر است؟

ج. اکنون می‌خواهیم دانش اولیه‌مان در مورد منصفانه بودن سکه را به تخمین اضافه کنیم. پارامتر این توزیع را با استفاده از روش MAP تخمین بزنید. احتمال اولیه را یک توزیع یکنواخت در نظر بگیرید.

د. با توجه به داده‌های فوق مقدار عددی پارامتر تخمین‌زده شده با روش MAP چقدر است؟

ه. نتیجه بندهای ب و د را با هم مقایسه کنید.

و. تخمین MAP پارامتر در حالتی که احتمال اولیه μ یک تابع توزیع گاوسی با میانگین $\frac{1}{2}$ و واریانس $\frac{1}{4}$ باشد به چه شکلی در می‌آید؟

ز. مقدار عددی این پارامتر را طبق تخمین محاسبه شده در قسمت و محاسبه نمایید.

۲- (اختیاری) سوال امتحانی میان ترم ۱۳۹۵

می‌خواهیم درجه حرارت T در داخل یک کوره را با استفاده از اندازه‌گیری‌های ثبت شده توسط n حسگر $\mathbf{x}=[x_1,x_2,\dots,x_n]$ کار گذاشته شده در نقاط مختلف آن تخمین بزنیم. خروجی هر حسگر بصورت مستقل بوسیله یک نویز افزودنی گوسی با میانگین صفر و واریانس σ_i^2 (برای حسگر i ام) آلوده شده است.

الف- تابع شباهت اندازه‌گیری‌ها ($\mathbf{x}$) بازای T=t داده شده را بدست آورید.

ب- تخمین ماکزیمم شباهت درجه حرارت ($\hat{T} = MLE(T|\mathbf{x})$) را محاسبه کنید.

۳- (اختیاری) یک متغیر تصادفی x از توزیع Erlang بصورت زیر تبعیت می‌کند:

$$p(x;\theta) = \theta^2 x e^{-\theta x} U(x)$$

که در آن $U(x)$ تابع پله واحد است. چنانچه N نمونه x_k $k=1,\dots,N$ در اختیار باشد، تخمین ماکزیمم شباهت θ را بدست آورید.

۴- (اختیاری) متغیر تصادفی x دارای توزیع نرمال $N(\mu,\sigma^2)$ با پارامتر مجهول μ می‌باشد. چنانچه بدانیم μ از توزیع ریلی بصورت زیر پیروی می‌کند:

$$p(\mu) = \frac{\mu e^{-(\mu^2/2\sigma_\mu^2)}}{\sigma_\mu^2}$$

نشان دهید که تخمین MAP پارامتر μ از رابطه زیر بدست می‌آید:

$$\hat{\mu}_{MAP} = \frac{Z}{2R}\left(1+\sqrt{1+\frac{4R}{Z^2}}\right)$$

که در آن:

$$Z = \frac{1}{\sigma^2}\sum_{k=1}^{N} x_k, \qquad R = \frac{N}{\sigma^2} + \frac{1}{\sigma_\mu^2}$$

۵- (اختیاری) مطلوب است محاسبه روابط 2.98، 2.99 و 2.100 مربوط به الگوریتم بیشینه سازی امید (EM) در کتاب تئودوریدیس (ویرایش چهارم ۲۰۰۹).

۶- (اختیاری) فرض کنید P احتمال قرارگیری k نمونه (از میان N نمونه) در یک ناحیه R از فضای ویژگی باشد. می‌دانیم k از توزیع دو-جمله‌ای بصورت زیر پیروی می‌کند:

$$P(k) = \frac{N!}{k!(N-k)!} P^k (1-P)^{N-k}$$

چنانچه آزمایش m بار تکرار شود و هر بار N نمونه‌ی تصادفی تولید شده و k_i تعداد نمونه‌های قرار گرفته در ناحیه R شمارش گردد ($i=1,\dots,m$)، در این صورت رابطه تخمین پارامتر P در توزیع دوجمله‌ای را بر اساس روش ماکزیمم شباهت بدست آورید. (راهنمایی: می-توانید از تابع شباهت لگاریتمی استفاده کنید).

۷- (اختیاری) مساله امتحان میان ترم بازشناسی آماری الگو سال ۱۳۸۷

چنانچه نمونه های $k=1,\dots,N$ $\mathbf{x}_k$ در فضای دوبعدی از یک توزیع گوسی با میانگین $\mathbf{m} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ و ماتریس کواریانس نامعلوم Σ بدست آمده باشند، مطلوب است محاسبه تخمین ماتریس کواریانس با استفاده از روش ماکزیمم شباهت:

$$\hat{\Sigma}_{ML} = ?$$

۸- (اختیاری) مساله امتحان میان ترم بازشناسی آماری الگو سال ۱۳۹۱

فرض کنید X در بازه $[\theta, 1]$ از توزیع یکنواخت به صورت زیر تبعیت می‌کند:

$$f_\theta(x) = \begin{cases} \frac{1}{1-\theta} & \theta < x < 1 \\ 0 & otherwise \end{cases}$$

اگر n نمونه x_1 تا x_n به صورت مستقل و با استفاده از توزیع بالا تولید شده باشند، تخمین ماکزیمم شباهت پارامتر θ را به دست آورید.

۹- (اختیاری) مساله امتحان میان ترم بازشناسی آماری الگو سال ۱۳۹۶

فرض کنید:

$$x_1, x_2, \dots, x_N \sim Poisson(\lambda)$$

چنانچه:

$$\lambda \sim Gamma(\alpha, \beta)$$

مطلوب است تخمین λ با استفاده از نمونه های داده شده. روابط توزیع پواسون و گاما بصورت زیر می باشد:

$$Poisson(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad Gamma(\lambda, \alpha, \beta) = \lambda^{\alpha-1} e^{-\lambda\beta}$$

۱۰- در فضای یک بعدی فرض کنید یک تابع چگالی احتمال بصورت مجموع وزن‌دار دو توزیع گوسی با واریانس واحد و میانگین های نامعلوم μ_1 و μ_2 بصورت زیر باشد:

$$p(x) = P_1 p(x|\mu_1) + P_2 p(x|\mu_2)$$

در رابطه فوق P_1 و P_2 وزنهای هریک از توزیع های گوسی هستند و معلوم فرض می شوند.

الف- چنانچه یک مجموعه نمونه بصورت $X=\{x_1,x_2,\dots,x_N\}$ در اختیار داشته باشیم که از توزیع $p(x)$ بدست آمده باشد، مطلوب است تعریف تابع شباهت برای تخمین μ_1 و μ_2.

ب- با بیشینه سازی تابع شباهت تخمین μ_1 و μ_2 را بدست آورید. ($\frac{dlnu(x)}{dx}=\frac{1}{u(x)}\frac{du(x)}{dx}$)

۱۱- (اختیاری) فرض کنید $X\sim Unif(0,\theta)$. با فرض داشتن نمونه‌های $\{X_1,\dots,X_n\}$ که i.i.d هستند، تخمین بیشینه درست‌نمایی پارامتر θ را بیابید.

۱۲- (اختیاری) (سوال امتحانی میان ترم ۱۳۹۸)

فرض کنید s[n] یک سیگنال معلوم (مثلا تابع پله واحد) و w[n] نویز گوسی با میانگین صفر و واریانس معلوم σ^2 باشد و داشته باشیم:

$$x[n]=As[n]+w[n] \qquad n=0,1,\dots,N-1$$

چنانچه پارامتر A مجهول باشد، مطلوب است تخمین بیشنه شباهت A با استفاده از نمونه های x[n].

۱۳- (اختیاری) (سوال امتحانی میان ترم ۱۳۹۹) اگر به تعداد n نمونه از یک توزیع احتمال به صورت $x_1,x_2,x_3,\dots,x_n$ داشته باشیم و این توزیع به صورت زیر باشد:

$$p(x|\theta)=\begin{cases}(\theta+1)x^{\theta} & for\ 0\le x\le 1\\ 0 & otherwise\end{cases}$$

همچنین اگر بدانیم که :

$$p(\theta)=e^{-\theta+1} \qquad for\ \ \theta>1$$

الف- تخمین MAP به چه صورت میشود ؟

ب- تخمین MLE برای این توزیع چگونه است ؟

۱۴- (اختیاری) فرض کنید توزیع احتمالی متغیر تصادفی گسسته X به صورت زیر باشد:

$$P(x=k)=\begin{cases}2-\sqrt{e} & k=0\\ \frac{1}{2^k k!} & k=1,2,\dots\end{cases}$$

الف) میانگین X را به دست آورید.

ب) واریانس X را به دست آورید.

۱۵- (میان ترم ۱۴۰۱) یک مساله دسته بندی دو کلاسه را در فضای یک بعدی در نظر بگیرید. چنانچه توابع چگالی احتمال کلاس ها بصورت زیر داده شده باشند:

$$p(x/c_1)=\begin{cases}0 & if\ x<0\\ \theta_1 e^{-\theta_1 x} & if\ x\ge 0\end{cases} \qquad p(x/c_2)=\begin{cases}0 & if\ x<0\\ \theta_2 e^{-\theta_2 x} & if\ x\ge 0\end{cases}$$

چنانچه $D_1=\{3,5\}$ مجموعه نمونه‌های موجود کلاس ۱ و $D_2=\{6,9,12\}$ مجموعه نمونه‌های کلاس ۲ باشد:

الف- تخمین بیشینه شباهت پارامترهای θ_1 و θ_2 را بدست آورید.

ب- با استفاده از دسته‌بند بیشینه شباهت (مقایسه توابع چگالی احتمال)، نواحی تصمیم گیری و مرزهای تصمیم گیری را بر اساس پارامترهای تخمین زده شده در بند قبل بدست آورید.

۱۶- (میان ترم ۱۴۰۲)

تابع چگالی احتمال زیر را در نظر بگیرید:

$p_\theta(x) = 2\theta x e^{-\theta x^2}$

که در آن پارامتر θ و x مقادیر حقیقی و مثبت هستند. فرض کنید تعداد N نمونه x_i از این توزیع در اختیار باشد. مطلوب است تخمین پارامتر θ با استفاده از روش بیشینه شباهت.

۱۷- (میان ترم ۱۴۰۳)

یک متغیر تصادفی x از یک توزیع بصورت زیر تبعیت می‌کند:

$$p(x;\theta) = \frac{\theta^3}{2} x^2 e^{-\theta x}, x > 0, \theta > 0$$

چنانچه N نمونه x_k $k=1,\dots,N$ در اختیار باشد، تخمین ماکزیمم شباهت θ را بدست آورید.

موفق باشید