دانشکده‌ی مهندسی کامپیوتر

آمار و احتمال مهندسی بهار ۱۳۹۹

تمرین سری هشتم (تخمین و تشخیص)

مدرس: نعیمه امیدوار موعد تحویل: سه‌شنبه ۳ تیر ۱۳۹۹

* سوالات ۶، ۷ و ۸ اختیاری هستند.

سؤال ۱ یک معیار مهمی که می‌تواند برای سنجیدن عملکرد تخمین به‌کار گرفته شود مقدار bias آن تخمین است. این عبارت را به‌صورت زیر تعریف می‌کنیم:

$$bias(\widehat{\theta}) = E[\widehat{\theta}] - \theta$$

($\widehat{\theta}$ تخمینگری برای پارامتر θ است.)
فرض کنید n متغیر تصادفی $x_1, x_2, ..., x_n$ داریم اگر امیدریاضی آن‌ها μ باشد برای واریانس آن از رابطه‌ی زیر استفاده می‌کنیم:

$$\sigma^2 = E[(X-\mu)^2]$$

دو رابطه‌ی زیر برای تخمین واریانس واریانس این n داده پیشنهاد می شود. این دو تخمین را بر اساس بایاس مقایسه کنید و بگویید کدام یک تخمینگر بهتری است.

$$S_1^2 = \frac{1}{n}\sum_1^n (x_i - \bar{x})^2$$

$$S_2^2 = \frac{1}{n-1}\sum_1^n (x_i - \bar{x})^2$$

سؤال ۲ فیزیکدانی ماهر یک آشکارساز ذرات ساخته‌است که از آن برای محاسبه‌ی کمیت‌های فیزیکی سیارات و ستاره‌ها استفاده می‌کند. فرض کنید در طی یک شبانه روز تعداد ذراتی که به این آشکارساز برخورد می‌کنند (Y) از یک توزیع احتمالی شرطی پواسونی نسبت به پارامتر x پیروی کنند. پارامتر x نامعلوم است و به صورت مقداری از متغیر تصادفی X که از توزیع نمایی با پارامتر μ تبعیت می‌کند مدلسازی می‌شود. مقدار μ مشخص است. بنابراین داریم:

$$f_X(x) = \begin{cases} \mu e^{-\mu x} & x \geq 0 \\ 0 & \text{O.W} \end{cases} \tag{۱}$$

$$P_{Y|X}(y|x) = \begin{cases} \frac{e^{-x}x^y}{y!} & y = 0, 1, 2, ... \\ 0 & \text{O.W} \end{cases} \tag{۲}$$

الف. با استفاده از تخمینگر MAP متغیر x را با استفاده از تعداد ذرات آشکارشده‌ی y تخمین بزنید.
ب. در این قسمت می‌خواهیم امیدریاضی شرطی متغیر تصادفی X را نسبت به تعداد ذرات آشکار شده‌ی y به‌دست آوریم. به عبارت بهتر با استفاده از تخمینگر MMSE این متغیر را تخمین بزنیم.

۱. نشان دهید توزیع احتمال پسین X نسبت به Y از رابطه‌ای به فرم زیر تبعیت می‌کند و پارامتر λ را نیز نسبت به پارامترهای مساله به‌دست آورید.

$$f_{X|Y}(x|y) = \frac{\lambda^{y+1}}{y!}x^y e^{-\lambda x}$$

$$Tip: \int_0^\infty \alpha^{y+1}x^y e^{-\alpha x}dx = y! \quad for\ any\ \alpha > 0$$

۲. با استفاده از نتیجه‌ی قسمت قبل امیدریاضی شرطی متغیر تصادفی X را نسبت به تعداد y ذرات مشاهده شده به‌دست آورید.

سؤال ۳ نمونه‌های $X_1, X_2, X_3, \ldots$ را درنظر بگیرید که از یک توزیع برنولی پیروی می‌کنند. احتمال موفقیت در این توزیع را q درنظر بگیرید که مقداری نامعلوم است. متغیر زمانی T_K را به صورت زیر تعریف می‌کنیم:

$$T_1 = Y_1 \quad , \quad T_k = Y_k - Y_{k-1} \quad k = 2, 3, \ldots$$

که متغیر Y_k زمان kمین موفقیت در این فرایند است. در این مساله سعی می‌کنیم q را با استفاده از نمونه‌های زمانی $t_1, t_2, t_3, \ldots$ تخمین بزنیم. توزیع پیشین (prior distribution) متغیر q را متغیرتصادفی Q درنظر بگیرید که از توزیع یکنواخت $[0, 1]$ پیروی می‌کند.
الف. تابع جرم احتمال (PMF) متغیر T_1 یعنی $P_{T_1}(t_1)$ را به‌دست آورید.
ب. تخمین MMSE متغیر q را با استفاده از اولین داده یعنی $T_1 = t$ به‌دست آورید.
تخمین MAP متغیر q را با استفاده از k داده‌ی اول یعنی $T_1 = t_1, T_2 = t_2, \ldots, T_k = t_k$ به‌دست آورید.
راهنمایی : از انتگرال زیر می‌توانید استفاده کنید.

$$\int_0^1 q^k (1-q)^m dq = \frac{k!m!}{(k+m+1)!}$$

سؤال ۴ فرض کنید توزیع توام دو متغیر تصادفی X و Y از رابطه‌ی زیر تبعیت کند:

$$f_{XY}(x, y) = \begin{cases} cxy & if\ 0 < x \leq 1\ ,\ 0 < y \leq 1 \\ 0 & \text{O.W} \end{cases} \tag{۳}$$

الف. مقدار ضریب نرمالسازی c را به‌دست آورید.
ب. تخمین MMSE متغیر تصادفی X را براساس داده‌ی مشاهده شده‌ی $Y = y$ به‌دست آورید.
پ. آیا تخمینی که در قسمت قبل به‌دست آوردید با حالتی که هیچ اطلاعی از داده‌ی $Y = y$ داشتیم متفاوت می‌بود یا نه؟ توضیح دهید.
ت. دو قسمت قبل را این بار برای تخمینگر MAP متغیر X انجام دهید.

سؤال ۵ در این سوال قصد داریم مساله‌ی رگرسیون خطی را با رویکرد Bayesian بررسی کنیم. فرض کنید n جفت داده‌ی (x_1, y_1) و (x_2, y_2) و ... (x_n, y_n) داریم. ساده ترین مدلی که به ذهنمان می‌رسد این است که یک رابطه‌ی خطی میان این داده‌ها برقرار است. یعنی Y یک تابع خطی از X است. به عبارت دیگر:

$$y_i = ax_i + b$$

در این مساله قصد داریم با رویکرد بیزین دو پارامتر a و b را تخمین بزنیم. در این صورت برای یک داده‌ی x' که قبلا مشاهده نشده می‌توانیم y متناظر را تخمین بزنیم. برای سادگی فرض کنید که x_i و y_i یک بعدی هستند.
معادله‌ی $y_i = ax_i + b$ را به شکل دیگری نیز می‌توانیم بنویسیم:

$$y_i = w^T X_i \quad , \quad w = \begin{bmatrix} a \\ b \end{bmatrix} \quad , \quad X_i = \begin{bmatrix} x_i \\ 1 \end{bmatrix}$$

در واقع یک سطر به x اضافه کردیم تا این رابطه را به‌صورت ضرب ماتریسی بنویسیم. سپس برای سادگی فرض کنید

$$P(y \mid X, w) \sim N(w^T X, \sigma^2)$$

و فرض کنید مقدار σ را می‌دانیم. هدف نهایی این است که توزیع پسین posterior distribution ماتریس w را حساب کنیم. به عنوان توزیع پیشین برای این متغیر فرض کنید

$$P(w) \sim N(\mu_0 = 0, \sigma_0^2) \propto exp(-\frac{w^2}{2\sigma_0^2})$$

الف. حال با استفاده از نکات بالا رابطه‌ی زیر را برای توزیع پسین w به‌دست آورید:

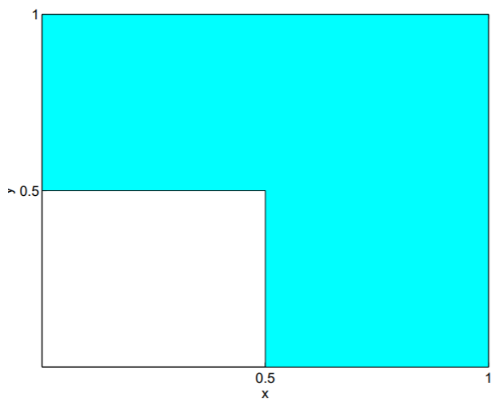
$$log(P(w \mid y, X)) \propto -\frac{1}{2\sigma^2} \sum (y_i - w^T x_i)^2 - \frac{1}{2\sigma_0^2} w^2$$

ب. حال می‌خواهیم رابطه‌ی بالا نسبت به پارامتر w یعنی a و b بیشینه شود. بنابراین با مشتق‌گیری نسبت به این پارامترها a و b را برحسب رابطه‌ای از x_iها و y_iها تخمین بزنید.

بخش عملی سوال:

یک فایل data.csv در کنار این فایل قرار گرفته‌است که شامل دو ستون X و Y است. با استفاده از زبان R یا Python یک تابع بنویسید که با دریافت این فایل به عنوان ورودی
الف. دو مقدار a و b را محاسبه کرده و آن‌ها را خروجی دهد.
ب. یک نمودار رسم کند که داده‌های x و y به‌صورت scatter روی آن رسم شده‌باشند و خط $y = ax + b$ نیز روی این داده‌ها رسم شده باشد تا بتوانیم دقت خروجی‌مان را با استفاده از نمودار بسنجیم.
(منظور از توان ۲ی بردار یا ماتریس درواقع جمع توان ۲ی تمام درایه‌های آن است همچنین برای سادگی فرض کنید $\sigma_0 = \sigma = 1$)

سؤال ۶ در شکل زیر نقاط تصادفی (X, Y) به صورت یکنواخت روی ناحیه آبی پخش شده‌اند.



(الف) تخمینگر MMSE ، $\hat{y}_{mmse}(x)$ را محاسبه کنید.

(ب) تخمینگر LMMSE ، $\hat{y}_{lmmse}(x)$ را محاسبه کنید.

سؤال ۷ فرض کنید تخمینگر زیر را برای پارامتر θ داریم که مبنی بر مشاهده یک نمونه (x_0) است.

$$\hat{\theta} = ax_0^2 + bx_0 + c$$

ضرایب a, b, c را چنان بیابید که خطای MSE مربوط به این تخمینگر کمینه شود. میدانیم x_0 از توزیع یکنواخت در بازه $[-0.5, 0.5]$ است و $\theta = cos(2\pi x_0)$، تخمینگر LMMSE را نیز برای θ بیابید. MSE کمینه هر دو تخمینگر را یافته و با هم مقایسه کنید.

سؤال ۸ فرض کنیم توزیع های زیر را می‌دانیم.

$$X \sim N(\mu, \tau^2) \quad , \quad Y|X = x \sim N(x, \sigma^2)$$

(الف) توزیع $X|Y = y$ را محاسبه کنید.

(ب) فرض کنید که ما n وقوع تصادفی $(Y_1, ..., Y_n)$ از توزیع $Y|X = x$ را مشاهده کرده‌ایم. حال توزیع $X|\bar{Y}$ را محاسبه کنید که در آن $\bar{Y}$ میانگین نمونه ($\bar{Y} = \frac{\sum_i Y_i}{n}$) است. (میتوانید از نتایج الف استفاده کنید.)

(پ) تخمینگر MAP و MMSE را برای X بر حسب مشاهده $\bar{Y}$ بیابید.

سؤال ۹ فرض کنید $X \sim Uniform(1,2)$ و همچنین میدانیم توزیع $Y|X=x$ نمایی با پارامتر $\lambda = \frac{1}{x}$ است.

(الف) تخمین LMMSE ، X بر حسب Y را بیابید. ($\hat{X}_L$)

(ب) خطای MSE مربوط به تخمین بخش الف را بیابید.

(پ) نشان دهید $E[(X - \hat{X}_L)Y] = 0$.

سؤال ۱۰ فرض کنید یک پارامتر تصادفی $\Theta \in R$ داریم که تابع PMF آن به صورت $P(\theta) = \frac{1}{2}(\delta(\theta-1) + \delta(\theta+1))$ است (δ بیانگر تابع دلتای دیراک ست که در صورتی که $x = 0$ باشد مقدار $\delta(x) = 1$ و در غیر این صورت مقدار صفر به خود میگیرد). فرض کنیم که ما خروجی نویزی Y را مشاهده کرده‌ایم.

$$Y = \Theta + W$$

که در آن $W \sim N(0, \sigma^2)$ است و مقدار σ^2 در آن معین است. تخمینگر MMSE و MAP را برای پارامتر Θ محاسبه کنید و بحث کنید تحت چه شرایطی این دو تخمین تقریبا یکی هستند.

موفق باشید