عنوان بخش

مسئله‌ی ۱. احتمال و آمار؟

الف

در دو خط، تفاوت کلی علم احتمال و آمار را توضیح دهید.

ب

علم احتمال، از روش استدلال ...... برای رسیدن از ..... به ..... استفاده میکند. علم آمار از روش استدلال .... برای رسیدن از ...... به ..... استفاده میکند.

ج

تفاوت های کلی دیدگاه frequentist و Bayesian را توضیح دهید

حل.

الف

علم احتمال ویژگی یک جامعه بزرگ را میداند و با استفاده از آن احتمال یک رویداد را پیش بینی میکند. علم آمار با استفاده از نمونه هایی از یکی جامعه، میخواهد ویژگی های آن جامعه را تخمین بزند.

ب

استنتاجی، کل، جز
استقرایی، جز، کل

ج

در شیوه فراوان گرایانه، ویژگی جامعه را یک مجهول غیر تصادفی میبیند که با نمونه گیری به تعداد زیاد با استفاده از قانون اعداد بزرگ به دست می آورد. در شیوه بیزی، آن ویژگی را یک متغیر تصادفی میبیند و در تلاش برای پیدا کردن توزیع آن است ▷

مسئله‌ی ۲. آمار یکنواخت

فرض کنید توزیع $uniform(a,b)$ داده شده است.

الف

۳ داده ی ۱۰ و ۸ و ۱۲/۵ از این توزیع انتخاب شده اند. maximum likelyhood را برای پارامترهای a و b تخمین بزنید

ب

اگر n داده از این توزیع به صورت $x_۱, x_۲, ..., x_n$ انتخاب شده باشند، maximum likelyhood را برای پارامترهای a و b تخمین بزنید

حل.

الف

تابع likelyhood برای این سه داده با توجه به pdf توزیع یونیفورم به شکل زیر است:

$$f(data|a,b) = \begin{cases} \frac{۱}{(b-a)^۳} & a \leqslant data \leqslant b \\ ۰ & o.w \end{cases}$$

پس برای اینکه این تابع مقدار بیشینه خود را بگیرد، مخرج این تابع باید کمینه شود. برای این کار a مینیمم داده ها و b ماکسیموم آنهاست
$a = ۸, b = ۱۲/۵$

ب

همین موضوع را تعمیم میدهیم و a مینیمم x_i ها و b ماکسیموم آنهاست ▷

مسئله‌ی ۳. تاس های نامتقارن

یک تاس k وجهی به شما داده است. همچنین یک مجموعه نمونه از نتایج پرتاب این تاس در اختیار داریم. تعداد آمدن هر وجه تاس در مجموعه m_k است. به صورتی که

$$\sum_{i=۱}^{i=k} m_i = N$$

احتمال آمدن هر وجه تاس نیز با μ_i نمایش داده می‌شود. همچنین $X(x) = k$ نشان دهنده مقدار متغیر تصادفی وجه تاس است.(به عنوان مثال $X(\omega) = ۱$ یعنی وجه اول آمده است) تعداد کل پرتاب ها انجام شده N تعداد می‌باشد.

الف

فرض کنید که تاس دو وجهی است یک تخمین گر بیشینه درست نمایی برای احتمال هر وجه پیدا کنید.(دقت کنید که راه حل شما باید حتما شامل maximum likelihood باشد و به اینکه از نظر فیزیکی آیا تاس ۲ وجهی ممکن است یا خیر نیز توجه نداشته باشید.)

ب

فرض کنید که k وجه داریم و مساله را به صورت کلی حل کنید. بدان معنا که یک تخمین گر بیشینه درست نمایی برای هر یک از μ_K بدست آورید. (دقت کنید که راه حل شما باید شامل maximum likelihood باشد و راه حل هایی نظیر استقرا نمره ای نخواهند گرفت.)

حل.

الف

$$\mu_2 = 1 - \mu_1$$

$$M = \{m_1, m_2\}$$

$$L(m1, m2; \mu_1, \mu_2) = \mu_1^{m_1}(\mu_2)^{m_2} \longrightarrow L(m1, m2; \mu_1) = \mu_1^{m_1}(1-\mu_1)^{m_2}$$

$$ln(L(M;\mu_1)) = m_1 ln(\mu_1) + m_2 ln(1-\mu_1)$$

$$\hat{\mu_1} = \underset{\mu_1}{\operatorname{argmax}}\, ln(L(M;\mu_1)) = \underset{\mu_1}{\operatorname{argmax}}\, m_1 ln(\mu_1) + m_2 ln(1-\mu_1)$$

$$\tfrac{\partial ln(L(M;\mu_1)}{\partial \mu_1} = \tfrac{m_1}{\mu_1} - \tfrac{m_2}{1-\mu_1} = 0 \longrightarrow$$

$$m_1 - \mu_1 m_1 = m_2 \mu_1 \longrightarrow \mu_1 = \tfrac{m_1}{m_1+m_2} \longrightarrow \mu_2 = \tfrac{m_2}{m_1+m_2} \checkmark$$

ب

$$M = \{m_1, m_2, ..., m_k\}$$

$$\theta = \mu_1, ..., \mu_k$$

$$L(M;\theta) = \textstyle\prod_{i=1}^{i=k} \mu_i^{m_i}$$

$$\textstyle\sum_{i=1}^{i=k} m_i = N$$

حال باید تابع بالا را ماکسیمم کنیم. دقت کنید که از قوانین احتمال می‌دانیم که مجموع احتمال تمام وجوه برابر ۱ است.

$$\textstyle\sum_{i=1}^{i=k} \mu_i = 1$$

بنابراین باید تابع likelihood را به صورت مقید ماکسیمم کنیم.

$$M = \{m_1, m_2, ..., m_k\}$$

$$\theta = \{\mu_1, ..., \mu_k\}$$

$$L(M;\theta) = \prod_{i=1}^{i=k} \mu_i^{m_i}$$

$$maximize\ L\ s.t\ \sum_{i=1}^{i=k} \mu_i = 1$$

با استفاده از ضرایب لانگراژ داریم.

$$maximize\ L\ s.t\ g(\theta) = \sum_{i=1}^{i=k} \mu_i = 1$$

$$ln(L(M;\theta)) = \sum_{i=1}^{i=k} m_i \mu_i$$
$$\mathcal{L}(ln(L(M;\theta)), g(\theta)) = \frac{\partial ln(L(M;\theta))}{\partial \mu_j} - \lambda \frac{g(\theta)-1}{\partial \mu_j} = 0 \longrightarrow$$

$$\frac{m_j}{\mu_j} = \lambda \longrightarrow m_j = \mu_j \lambda$$

$$\sum_{i=i}^{i=k} m_j = N \longrightarrow \sum_{i=i}^{i=k} \lambda \mu_j = N \longrightarrow \lambda \sum_{i=i}^{i=k} \mu_j = N \longrightarrow$$

$$\lambda = N \longrightarrow \mu_i = \frac{m_i}{N} \checkmark$$

$\triangleright$

مسئله‌ی ۴. بازه های شکست

یک متغیر برنولی با احتمال موفقیت p را در نظر بگیرید که آنرا نمی‌دانیم. حال متغیر T را به صورت زیر برای هر موفقیت تعریف میکنیم.
T_i برابر تعداد شکست های بین موفقیت $T_{i-۱}$ و T_i است.(با احتساب موفقیت K ام)

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

که متغیر Y_K برابر شماره آزمایشی است که K امین موفقیت را در پی داشته است.
دقت کنید که تعداد کل آزمایش ها به شما داده نشده است بلکه تنها مجموعه T به شما داده است که برابر.

$$T = \{T_۱, T_۲, ..., T_k\}$$

است.

الف

حال یک تخمین گر بیشینه درست نمایی برای پارامتر P بدست‌آورید.
(دقت کنید که راه حل شما باید به صورت maximum likelihood باشد)

ب

با فرض دانستن p به نظر شما توزیع T_i به چه توزیع ای شبیه است($PMF_{T|p}$)؟ میانگین و واریانس این توزیع را بدست آورید.. سپس با استفاده از توزیع T و با فرض ندانستن p و همچنین هیچ گونه پیشفرض از داده های بدست آمده (یعنی فرض کنید که هیچ داده ای نیز به شما داده نشده است.) تابع PMF_T را بدست آورید.

ج

نشان دهید که رابطه زیر برقرار است. (p^* برابر مقدار واقعی پارامتر توزیع است.)

$$\forall \epsilon > ۰ \ : \ \lim_{k \to \infty} P(|\frac{۱}{\hat{P}} - \frac{۱}{p^*}| > \epsilon) = ۰$$

د

فرض کنید که $p^* > ۰/۵$ است. حال یک کران پایین برای K بیابید به صورتی که رابطه زیر برقرار باشد. (از نامساوی Chebyshev استفاده کنید.)

$$P(|\frac{۱}{\hat{P}} - \frac{۱}{p^*}| \leqslant ۰/۱) \geqslant ۰/۹۵$$

حل.

الف

تابع likelihood را می‌نویسیم. دقت کنید که تعداد کل پرتاب ها برابر جمع t_i ها است.

$$L(T;p) = p^k(۱-p)^{(\sum_{t_i \in T} t_i) - k}$$

$$ln(L(T;p)) = k \times ln(p) + ((\sum_{t_i \in T} t_i) - k) \times ln(۱-p)$$

$$\hat{p} = \underset{p}{\operatorname{argmax}}(ln(L(T;p)))$$

$$\frac{\partial ln(L(T;p))}{\partial p} = \frac{k}{p} - \frac{(\sum_{t_i \in T} t_i) - k}{۱-p} = ۰ \longrightarrow$$

$$\hat{p} = \frac{k}{\sum_{t_i \in T} t_i}$$

ب

همانطور که مشخص است T ها یک توزیع هندسی هستند زیرا تعداد ازمایشات تا رسیدن به اولین موفقیت را نشان می‌دهند. بنابراین میدانیم که :

$$\mu = \frac{۱}{p}$$

$$var = \frac{۱-p}{p^۲}$$

$$PMF_{T|p}(t) = (۱-p)^{t-۱}p \qquad t = ۱, ۲, ...$$

حال نمیدانیم که p چند است. بنابراین باید روی کل حالات توزیع حاشیه ای بنویسیم.

$$0 \leqslant p \leqslant 1$$

$$PMF_T(t) = \int_0^1 PMF_{T|p}dp = \int_0^1 (1-p)^{t-1}p\, dp = [-\frac{(1-p)^t(tp+1)}{t(t+1)}]_0^1 = \frac{1}{t(t+1)} \qquad t = 1, 2, ...$$

ج

از آنجا که مقدار $\frac{1}{\hat{P}}$ برابر مقدار بدست آمده برای میانگین داده ها است و مطابق با قانون اعداد بزرگ (دقت کنید که هر T_i یک iid است.) مقدار میانگین یک متغیر تصادفی با میانگین بدست آمده از داده های پایدار است و بایاس ندارد به عبارتی:

$$\forall \epsilon > 0 \ : \ \lim_{k\to\infty} P(|\frac{\sum_{t_i \in T} t_i}{k} - E(T_1)| > \epsilon) = 0$$

د

$$P(|\frac{\sum_{t_i\in T} t_i}{k} - E(T_1)| > \epsilon) \leqslant \frac{var(T_1)}{k\epsilon^2}$$

$$P(|\frac{\sum_{t_i\in T} t_i}{k} - \frac{1}{p^*}| \leqslant 0/1) = 1 - P(|\frac{\sum_{t_i\in T} t_i}{k} - E(T_1)| > 0/1) \geqslant 1 - \frac{var(T_1)}{k\times 0/01}$$

$$1 - \frac{var(T_1)}{k\times 0/01} \geqslant 0/95 \longrightarrow \frac{var(T_1)}{k\times 0/01} \leqslant 0/05$$

$$k \geqslant 2000(\frac{1-p}{p^2})$$

حداکثر عبارت $\frac{1-p}{p^2}$ برابر ۲ است. (دقت کنید که p بین $\frac{1}{2}$ و یک است. بنابراین:

$$k \geqslant 4000$$

▷

مسئله‌ی ۵. چزنوبیل

در چرنوبیل یک منبع در هر زمانی که اندازه گیری می شود K فوتون از خود ساطع می کند . ما فرض می کنیم که که K دارای توزیع زیر است :

$$p_K(k;\theta) = c(\theta)e^{-k\theta} \qquad k = 0, 1, 2, ...$$

بطوری که θ معکوس دمای منبع است و $c(\theta)$ ضریب نرمالیزاسیون است . همچنین فرض میکنیم که فوتون های ساطع شده از منبع هر بار مستقل از یکدیگر هستند . ما می خواهیم دمای منبع را با اندازه گیری های پی در پی فوتون های ساطع شده تخمین بزنیم .

الف

ضریب نرمالیزاسیون $c(\theta)$ را بیابید.

ب

تخمین ML را برای دمای منبع $\psi = \frac{۱}{\theta}$ را بر اساس $K_۱, K_۲, ...$ که **تعداد** فوتون های ساطع شده پس از n بار اندازه گیری است را بیابید.

حل.

الف

$$۱ = \sum_{k=۰}^{\infty} c(\theta)e^{-k\theta} = \frac{c(\theta)}{۱-e^{-\theta}}$$
$$c(\theta) = ۱ - e^{-\theta}$$

ب

$$L(\theta) = (۱ - e^{-\theta})^n \times e^{-\sum K_i \times \theta}$$
$$LL(\theta) = nln(۱ - e^{-\theta}) - \sum K_i \times \theta$$
$$\frac{dLL(\theta)}{d\theta} = \frac{ne^{-\theta}}{۱-e^{-\theta}} - \sum K_i = ۰$$
$$e^{-\theta} = \frac{\sum K_i}{n+\sum K_i} \implies \hat{\theta_n} = ln(۱ + \frac{n}{\sum K_i})$$
$$\implies \hat{\psi_n} = \frac{۱}{ln(۱+\frac{n}{\sum K_i})}$$

▷

مسئله‌ی ۶. بازی های شرطی

شما به عنوان یک آمار دان در شرکتی مشغول به کار شده اید. صاحبان کازینو می‌خواهند بدانند که برد افراد در بازی های شرطی شان به چه صورت است . شما در بررسی میزان برد افراد در یک بازی شرطی که توسط کازینو ها کنار گذاشته شده است به متغیر های $X_۱, X_۲, X_۳, ...$ که iid باتوزیع یکنواخت $U(-\theta, \theta)$ است رسیدید. ($\theta > ۰$)

الف

تخمین بیشینه درست نمایی $\hat{\Theta}$ را بیابید.

ب

توزیع دقیق $\hat{\Theta}$ را بدست آورید.

ج

آیا تخمین گر بدست آمده بدون بایاس است؟ در صورت وجود بایاس یک تخمین گر بدون بایاس ارائه دهید.

حل.

الف

$$L(\theta) = \frac{۱}{(۲\theta)^n} \qquad \theta \geqslant |X_۱|, |X_۲|, |X_۳|, ...$$
$$LL(\theta) = -n \times ln(۲\theta) \qquad \theta \geqslant |X_۱|, |X_۲|, |X_۳|, ...$$
$$\hat{\Theta} = max(|X_۱|, |X_۲|, |X_۳|, ...)$$

ب

$$F_{\hat{\Theta}}(a) = P(\hat{\Theta} < a) = P(max(|X_۱|, |X_۲|, |X_۳|, ...) < a)$$
$$\implies \quad F_{\hat{\theta}}(a) = P(|X_۱| < a)P(|X_۲| < a)...$$
$$F_{\hat{\Theta}}(a) = P(-a < X_۱ < a)P(-a < X_۲ < a)...$$
$$F_{\hat{\Theta}}(a) = (\frac{a}{\theta})^n$$
$$f_{\hat{\Theta}}(a) = n\theta^{-n}a^{n-۱} \qquad a \in [۰, \theta]$$

ج

امید ریاضی تخمین را حساب می کنیم

$$E[\hat{\Theta}] = \int_۰^\theta \alpha \times n\theta^{-n}\alpha^{n-۱}d\alpha = \frac{n}{n+۱}\theta$$

مشخص است که تخمینگر بایاس دارد اما بطور حدی در n های بزرگ بطور مایل بدون بایاس است . تخمینگر جایگزین بدون بایاس برای Θ بصورت زیر است :

$$\hat{\Theta} = \frac{n+۱}{n} \times max(|X_۱|, |X_۲|, |X_۳|, ...)$$

$\triangleright$

مسئله‌ی ۷. وقتی پارامتر نداریم!

تا بدین جا با تخمین هایی آشنا شدید که یک پارامتر تابع را تخمین می‌زند حال می‌خواهیم با یک تخمین دیگری آشنا شویم که به صورت نقطه ای توزیع را تخمین می‌زند.
فرض کنید یک تابع Q به شما داده است. همچنین یک متغیر h_q نیز به شما داده است که دارای ویژگی های زیر هستند:

$$\begin{cases} \forall u \ : \ ۰ \leqslant Q(u) \\ \\ \int_R Q(x)dx = ۱ \\ \\ \lim\limits_{N \longrightarrow \infty} Q(x_i) = \delta(x_i) \\ \\ \lim\limits_{N \longrightarrow \infty} \frac{۱}{h_q} Q(\frac{a-x_i}{h_q}) = \delta(a - x_i) \end{cases}$$

$$\begin{cases} u < \frac{۱}{۲} \ : \ Q(u) = ۱ \\ \\ o.w \ : \ Q(u) = ۰ \end{cases}$$

حال فرض کنید که مقدار تابع را به صورت زیر تخمین می‌زنیم. (و تعداد N داده به ما داده شده است.)

$$f(\hat{a}) = \frac{۱}{N*h_q} \sum_{q=۱}^{N} Q(\frac{a-x_q}{h_q})$$

در حقیقت تعداد داده هایی که در فاصله ای کمتر از $\frac{h_q}{۲}$ به تعداد کل داده هارا برای این تخمین گر استفاده می‌کنیم.(البته تقسیم بر طول بازه نیز می‌کنیم که یک تابع چگالی احتمال بدست آید.) فرض کنید که تابع $f^*(a)$ به شما داده شده است.

الف

ثابت کنید که این تخمین گر برای تابع بایاس ندارد. (دقت کنید که کانولوشن یک تابع با تابع دلتای دیراک برابر خود تابع است.)

حل. امید ریاضی تابع را بدست می‌اوریم.

الف

$$E(f(\hat{a})) = E(\frac{۱}{h_q} \sum Q(\frac{a-x_i}{h_q}))$$

$$\lim_{N \longrightarrow \infty} \int_R (\frac{۱}{N*h_q} \sum Q(\frac{a-x}{h_q}))f^*(x)dx =$$

$$\lim_{N \longrightarrow \infty} \frac{۱}{N} \sum \int_R (\frac{۱}{h_q} Q(\frac{a-x}{h_q}))f^*(x)dx =$$

$$\frac{۱}{N*h_q} \sum \int_R \delta(a-x)f^*(x)dx = \frac{۱}{N} * N * f^*(a) = f^*(a)\checkmark$$

امید ریاضی تابع با خود تابع اصلی در حد بی‌نهایت یکی است بنابراین تخمین گیر بدون بایاس است. (دقت کنید که مقدار انتگرال کانولوشن یک تابع با تابع دلتای دیراک است که برابر خود تابع است.) ◁

نکات مهم

- بخش تئوری را در قالب یک فایل pdf با اسم HW#_[STD-Num] آپلود کنید.
- ددلاین تمرین ساعت ۲۳:۵۹ روز ... می‌باشد.
-

موفق باشید :)