دانشکده‌ی مهندسی کامپیوتر

آمار و احتمال مهندسی

تمرین سری چهارم

مدرس: دکتر مطهری **مهلت: ۹ دی ساعت ۲۳:۵۵**

سوال ۱

توزیع $p(\theta|\alpha,\tau)$ را که $(\alpha > 1, \tau > 0)$، به صورت زیر تعریف می‌کنیم:

$$p(\theta|\alpha,\tau) \sim \begin{cases} \theta^{-(\alpha+1)}, & \theta \geq \tau \\ 0, & \text{otherwise} \end{cases}$$

الف

مقدار ضریب نرمال‌سازی این توزیع را بر حسب α و τ بیابید.

ب

توزیعی مانند $q(x|\theta)$ بیابید که توزیع q برای p، توزیع conjugate باشد، یعنی:

$$p(\theta|\alpha',\tau') = Prob(\theta|x) \sim q(x|\theta)p(\theta|\alpha,\beta)$$

سوال ۲

فرض کنید $X_1, X_2, \ldots, X_n$ متغیرهای تصادفی با توزیع $X_i \underset{i.i.d}{\sim} Unif(0,\theta)$ باشند.

الف

تخمین‌گر درستنمایی بیشینه [1]، $\hat{\theta}_{ML}$ را بیابید.

ب

فرض کنید $\hat{\theta}_{UB} = \alpha\hat{\theta}_{ML}$، مقدار α چقدر باشد تا تخمین‌گر $\hat{\theta}_{UB}$ نااریب [2] شود؟

---

[1] Maximum Likelihood Estimator

[2] Unbiased

ج

واریانس تخمین‌گر $\hat{\theta}_{UB}$ و کران کرامر-راو [۳] را حساب کنید. بعد از محاسبه، می‌بینید که واریانس $\hat{\theta}_{UB}$ کمتر از کران کرامر-راو است. علت این مسئله چیست؟

سوال ۳

فرض کنید $X_1, X_2, \dots, X_n$ متغیرهای تصادفی با توزیع $X_i \underset{i.i.d}{\sim} Poisson(\lambda)$ باشند. اگر از روش گشتاورها [۴] برای تخمین λ استفاده کنیم، تخمین‌گرهای $\hat{\lambda}_1$ و $\hat{\lambda}_2$ که به ترتیب تخمین‌گر با استفاده از برابری با گشتاور اول و تخمینگر با استفاده از برابری با گشتاور دوم هستند را به دست آورید.

سوال ۴

داده‌های $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ را در نظر بگیرید. می‌دانیم که $X_i \underset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$ و رابطه‌ی بین X_i و Y_i به صورت زیر مدل می‌شود:

$$Y_i = (w_1 X_i + w_0) + Z_i$$

در فرمول بالا $Z_j \underset{i.i.d}{\sim} \mathcal{N}(0, 1)$ هستند. هم‌چنین تمامی X_iها و Z_j ها از یکدیگر مستقل هستند.

به صورتی دیگر، می‌توانیم بگوییم که رابطه‌ی X و Y خطی است و Z مانند یک نویز گاوسی است که در هنگام نمونه‌گیری اضافه می‌شود. تخمین‌گر کمترین مربعات را به صورت زیر تعریف می‌کنیم:

$$\hat{w}_{LMS} = \arg\min_{w} \frac{1}{n} \sum_{i=1}^{n} (Y_i - (w_1 X_i + w_0))^2$$

الف

ثابت کنید که $\hat{w}_{LMS} = \hat{w}_{ML}$ که در اینجا $\hat{w}_{ML}$ نشان‌دهنده‌ی تخمین‌گر درست‌نمایی بیشینه است (راهنمایی: توزیع $P(Y|X, w_0, w_1)$ را بنویسید).

ب

اگر

$$\hat{w}_{RLMS} = \arg\min_{w} \frac{1}{n} \sum_{i=1}^{n} (Y_i - (w_1 * x_i + w_0))^2 + \frac{\lambda}{n} \|w\|_2^2$$

در این صورت به ازای چه مقداری از λ، $\hat{w}_{RLMS} = \hat{w}_{MAP}$ است که $\hat{w}_{MAP}$ نشان‌دهنده‌ی تخمین‌گر احتمال پسین بیشینه [۵] به ازای توزیع پیشین [۶] $P(w_0, w_1) = \mathcal{N}(w_0|0, 1)\mathcal{N}(w_1|0, 1)$ است (راهنمایی: توزیع $P(w_0, w_1|X, Y)$ را بر حسب $P(Y|X, w_0, w_1)$ و $P(w_0, w_1)$ بنویسید).

---

[۳] Cramér–Rao
[۴] Method of moments
[۵] Maximum A Posteriori Estimator
[۶] Prior Distribution

سوال ۵

فرض کنید $X_1, X_2, \dots, X_n$ متغیرهای تصادفی با توزیع $X_i \underset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$ باشند و توزیع پیشین پارامتر σ^2 برابر $\sigma^2 \sim \text{InverseGamma}(2, 1)$ است (ویژگی‌های توضیح معکوس-گاما را می‌توانید در صفحه ویکیپدیا [7] آن بخوانید).

الف)

توزیع پسین $P(\sigma^2 | X_1, X_2, \dots, X_n)$ را به دست آورید.

ب)

اگر تابع فاصله بر روی فضای پارامترها $L(\sigma_1^2, \sigma_2^2) = (\sigma_1^2 - \sigma_2^2)^2$ باشد، تخمین‌گر بیز را محاسبه کنید.

ج)

حال اگر داده‌ی جدید X_{n+1} را بگیریم، توزیع آن چیست؟ در واقع، توزیع $Prob(X_{n+1} | X_1, X_2, \dots, X_n)$ (به این توزیع، Predictive Distribution می‌گویند) را به دست آورید. این توزیع را نمی‌توان به صورت بسته محاسبه کرد و نوشتن آن به صورت یک انتگرال کافی است.

[7] https://en.wikipedia.org/wiki/Inverse-gamma_distribution