آمار و احتمال مهندسی

نیم‌سال دوم ۱۴۰۳-۱۴۰۲

مدرس: دکتر امیر نجفی

دانشکده‌ی مهندسی کامپیوتر

تمرین سری پنجم — زمان تحویل: ۴ خرداد

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص‌شده است.
- هم‌کاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

بارم‌بندی

بارم سوالات به شکل زیر است: (مجموعا ۱۰۰ نمره)

- سوالات ۱ و ۵: هر کدام ۱۵ نمره
- سوالات ۲، ۳، ۴، ۶ و ۸: هر کدام ۱۰ نمره
- سوال ۷: ۲۰ نمره

مسئله‌ی ۱. (برد و باخت)

در یک بازی برای هر بازیکن احتمال برد و باخت در هر دور با هم برابر و از دورهای دیگر مستقل است. یک بازیکن با هر برد یک امتیاز مثبت و با هر باخت یک امتیاز منفی دریافت میکند.

الف) برای متغیر تصادفی دلخواه X با استفاده از تابع مولد گشتاور [1] و نامساوی مارکوف [2] به ازای $t \geqslant ۰$ یک کران برای $\mathbb{P}(X \geqslant a)$ بدست آورید.

ب) تابع مولد گشتاورِ امتیازی را که یک بازیکن در هر دور کسب می‌کند، بدست آورید؛ سپس با استفاده از آن، یک کران برای تابع مولد گشتاور امتیاز کل یک بازیکن بعد از n دور بازی بدست آورید.

(راهنمایی: از بسط تابع e^x و نامساوی $(۲n)! \geqslant ۲^n n!$ استفاده کنید.)

ج) برای احتمال اینکه امتیاز یک بازیکن بعد از n دور بازی حداقل a باشد، یک کران بیابید و سپس این کران را در حالت $n = ۱۰۰, a = ۶۰$ بدست آورید.

مسئله‌ی ۲. (تخمین پارامتر)

متغیر تصادفی گسسته X را با تابع جرم احتمال زیر و شرط $\theta \in [۰, ۱]$ در نظر بگیرید:

$$P(X) = \begin{cases} \frac{۲\theta}{۳} & X = ۰ \\ \frac{\theta}{۳} & X = ۱ \\ \frac{۲-۲\theta}{۳} & X = ۲ \\ \frac{۱-\theta}{۳} & X = ۳ \end{cases}$$

۱۰ مشاهده مستقل از این توزیع به دست آمده‌اند:

$$(۱, ۲, ۰, ۱, ۲, ۳, ۱, ۲, ۰, ۳)$$

با توجه به این مشاهدات، تخمین بیشینه درست‌نمایی[3] را برای θ به دست آورید.

مسئله‌ی ۳. (تخمین‌گر سازگار)

اگر $X_۱, \cdots, X_n$ یک نمونه تصادفی از $\text{Bernoulli}(p)$ باشد، ثابت کنید که $\bar{X} = \frac{\sum_{i=۱}^{n} X_i}{n}$ یک تخمین‌گر سازگار برای p است.

[1] Moment generating function
[2] Markov's inequality
[3] Maximum Likelihood Estimation

مسئله‌ی ۴. (سکه)

سارا یک سکه را سه بار پرتاب می‌کند. سکه هر سه بار خط می‌آید. سپس او سکه را به شادی می‌دهد. شادی تا زمانی که اولین شیر را ببیند، سکه را پرتاپ می‌کند. در مجموع سکه ۴ بار پرتاب می‌شود. اگر θ را احتمال شیر آمدن سکه در نظر بگیریم، تخمین‌گر بیشینه درست‌نمایی را برای θ محاسبه کنید.

مسئله‌ی ۵. (میانگین مربعات)

متغیرهای تصادفی X، Y و W را در نظر بگیرید. فرض کنید داریم:

$$X \sim \mathcal{N}(0, 1), \qquad W \sim \mathcal{N}(0, 1), \qquad Y = X + W$$

اگر بدانیم که W مستقل از X است،

الف) کمینه میانگین مربعات خطا [4] تخمین‌گر X را به شرط داشتن Y بیابید. در بخش بعدی سوال، این مقدار را با $\hat{X}_M$ نشان خواهیم داد.

ب) میانگین مربعات خطا [5] را برای این تخمین‌گر محاسبه کنید. (راهنمایی: $\text{MSE} = \mathbb{E}[(X - \hat{X}_M)^2]$)

ج) فرض کنید $\tilde{X}$ نشان دهنده‌ی خطای بین مقدار واقعی X و تخمینگر MMSE آن، $\hat{X}_M$ باشد. به عبارت دیگر:

$$\tilde{X} = X - \hat{X}_M$$

نشان دهید رابطه‌ی $\mathbb{E}[X^2] = \mathbb{E}[\hat{X}_M^2] + \mathbb{E}[\tilde{X}^2]$ برقرار است.

مسئله‌ی ۶. (برنولی)

یک متغیر تصادفی برنولی با پارامتر p را در نظر بگیرید. p نامشخص است؛ اما می‌دانیم $p \in [\frac{1}{10}, \frac{1}{5}]$. فرض کنید که می‌توانیم این آزمایش را به تعداد دلخواه انجام دهیم و آزمایش‌ها دو به دو مستقل از یکدیگر هستند. تخمین‌گر زیر را برای پارامتر p در نظر بگیرید:

$$\hat{p} = \frac{\text{تعداد آزمایش‌های موفق}}{\text{تعداد کل آزمایش‌ها}}$$

حداقل چند آزمایش باید انجام دهیم تا مطمئن شویم که انحراف معیار تخمین‌گر مورد نظر کمتر از $\frac{1}{100}$ خواهد بود؟

---

[4] MMSE
[5] MSE

مسئله‌ی ۷. (آهن و بتن)

X و Y دو متغیر تصادفی مستقل هستند که به ترتیب مقاومت تیرهای آهنی و ستون‌های بتنی را نشان می‌دهند. داده‌های زیر پس از مشاهدات مختلف حاصل شده است:

$$X: ۵/۹,\ ۷/۲,\ ۷/۳,\ ۶/۳,\ ۸/۱,\ ۶/۸,\ ۷,\ ۷/۶,\ ۶/۸,\ ۶/۵,$$
$$۷,\ ۶/۳,\ ۷/۹,\ ۹,\ ۸/۲,\ ۸/۷,\ ۷/۸,\ ۹/۷,\ ۷/۴,\ ۷/۷,$$
$$۹/۷,\ ۷/۸,\ ۷/۷,\ ۱۱/۶,\ ۱۱/۳,\ ۱۱/۸,\ ۱۰/۷$$

$$Y: ۶/۱,\ ۵/۸,\ ۷/۸,\ ۷/۱,\ ۷/۲,\ ۹/۲,\ ۶/۶,\ ۸/۳,\ ۷,\ ۸/۳,\ ۷/۸,$$
$$۸/۱,\ ۷/۴,\ ۸/۵,\ ۸/۹,\ ۹/۸,\ ۹/۷,\ ۱۴/۱,\ ۱۲/۶,\ ۱۱/۲$$

حال فرض کنید داریم: $\mathbb{E}[X] = \mu_۱$، $\mathrm{Var}[X] = \sigma_۱^۲$، $\mathbb{E}[Y] = \mu_۲$ و $\mathrm{Var}[Y] = \sigma_۲^۲$.

الف. نشان دهید که $\bar{X} - \bar{Y}$ یک تخمین‌گر نااریب برای $\mu_۱ - \mu_۲$ است و آن را برای داده‌های بدست آمده محاسبه کنید.

ب. واریانس و انحراف معیار تخمین‌گر بخش (الف) را بدست آورده و مقدار انحراف معیار را برای داده‌های بدست آمده محاسبه کنید.

ج. تخمینی از نسبت $\frac{\sigma_۱}{\sigma_۲}$ بدست آورید.

د. یک تیرآهن با مقاومت X و یک ستون بتنی با مقاومت Y را به طور تصادفی انتخاب می‌کنیم. تخمینی برای واریانس اختلاف Y از X ($X - Y$) بدست آورید.

مسئله‌ی ۸. (میانگین یا میانه)

فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی احتمال زیر است:

$$f(x) = \frac{\lambda}{۲} \exp(-\lambda|x-\mu|)\,, \qquad -\infty < x < \infty$$

پارامتر های λ و μ نامعلوم هستند. امید ریاضی متغیر تصادفی X برابر با μ و توزیع آن نسبت به $x = \mu$ متقارن است؛ بنابراین میانه نیز همان μ می‌باشد.

ما یک نمونه تصادفی $X_۱, \cdots, X_n$ به بزرگی n داریم. تعیین کنید که میانگین نمونه، تخمین‌گر بهتری برای μ است یا میانه نمونه؟ توجه کنید در اینجا منظور از بهتر بودن، میزان واریانس کمتر است.