به نام خدا

دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس مبانی یادگیری آماری

استاد نیک‌آبادی

تمرین دوم

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

قسمت تئوری

توجه: برای حل سوالات قسمت تئوری از پاسخ‌نامه موجود در کورسز و گیت‌هاب https://github.com/telmo-correa/all-of-statistics الهام گرفته شده است.

سوال ۱

با توجه به آنکه تنها لازم است تا دو قسمت مربوط به همگرایی را ثابت کنیم، می‌توان از دو رابطه مربوط به امیدریاضی و واریانس در تئوری استفاده کرد:

$$MSE = bias^2\left(\widehat{F}_n(x)\right) + \mathbb{V}\left(\widehat{F}_n(x)\right) = \mathbb{E}\left(\widehat{F}_n(x) - \mathbb{E}\left(\widehat{F}_n(x)\right)\right)^2 + \mathbb{V}\left(\widehat{F}_n(x)\right)$$

$$= \mathbb{E}\left(\widehat{F}_n(x) - F(x)\right)^2 + \frac{F(x)\big(1-F(x)\big)}{n} = \frac{F(x)\big(1-F(x)\big)}{n}$$

مقدار $F(x)$ قطعا محدود است، مقدار $\big(1-F(x)\big)$ هم محدود است ولی مخرج از n است و وقتی به بی‌نهایت میل کند حاصل کل کسر یعنی MSE به صفر میل می‌کند.

$$\mathbb{V}\left(\widehat{F}_n(x)\right) = \frac{F(x)\big(1-F(x)\big)}{n}$$

$$\mathbb{E}\left(\widehat{F}_n(x)\right) = F(x)$$

به طور مشابه داریم:

$$\mathbb{V}\left(\widehat{F}_n(x)\right) \to 0 \Rightarrow se\left(\widehat{F}_n(x)\right) \to 0$$

و همچنین:

$$MSE \to 0 \Rightarrow bias \to 0$$

طبق قضیه 6.10 اگر با $n \to \infty$ داشته باشیم $bias \to 0$ و $se \to 0$ می‌توان نتیجه گرفت که $\widehat{F}_n(x)$ سازگار (consistence) است و در نتیجه:

$$\widehat{F}_n(x) \xrightarrow{P} F(x)$$

سوال ۲

مطابق با CLT می‌دانیم به صورت تقریبی رابطه زیر برقرار است:

$$\widehat{F_n}(x) \rightsquigarrow N\left(\mathbb{E}\left(\widehat{F_n}(x)\right), \frac{\mathbb{V}\left(\widehat{F_n}(x)\right)}{n}\right)$$

مطابق تئوری 7.3 که در سوال ۱ بود خواهیم داشت:

$$\widehat{F_n}(x) \rightsquigarrow N\left(F(x), \frac{F(x)\left(1 - F(x)\right)}{n^2}\right)$$

طبیعتا خواهیم داشت:

$$n \to \infty: \widehat{F_n}(x) \rightsquigarrow N(F(x), 0) \Rightarrow \widehat{F_n}(x) \rightsquigarrow F(x)$$

سوال ۳

$$f(x|\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}$$

تخمین با استفاده از مومان‌ها:

$$\alpha_1\left(\hat{\theta}\right) = \widehat{\alpha_1}$$

$$\alpha_j(\theta) = \mathbb{E}_\theta\left(X^j\right)$$

$$\widehat{\alpha_J} = \frac{1}{n}\sum_{i=1}^{n} X_i^j$$

$$\alpha_1\left(\hat{\theta}\right) = \mathbb{E}_\theta(X) = \hat{\lambda}$$

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

تخمین با استفاده از بیشینه درست‌نمایی:

$$\mathcal{L}(\theta) = f(X^n|\theta) = \prod_{i=1}^{n} f(X_i|\theta) = \prod_{i=1}^{n} \frac{\lambda^{X_i}}{X_i!} e^{-\lambda} \propto \prod_{i=1}^{n} \lambda^{X_i} e^{-\lambda}$$

$$\ell(\theta) = \log\big(\mathcal{L}(\theta)\big) \propto \sum_{i=1}^{n} \log\big(\lambda^{X_i} e^{-\lambda}\big) = \sum_{i=1}^{n} (X_i \log(\lambda) - \lambda)$$

$$= \left(\log(\lambda) * \sum_{i=1}^{n} X_i\right) - n\lambda$$

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0 \rightarrow \frac{1}{\lambda}\sum_{i=1}^{n} X_i = n \rightarrow \lambda = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

امتیاز فیشر:

$$\log f(X;\lambda) = x\log(\lambda) - \log(x!) - \lambda$$

$$s(X;\lambda) = \frac{\partial \log f(X;\lambda)}{\partial \lambda} = \frac{x}{\lambda} - 1$$

$$\frac{\partial^2 \log f(X_i;\lambda)}{\partial \lambda^2} = -\frac{x}{\lambda^2}$$

$$I(\lambda) = -\mathbb{E}\left(\frac{\partial^2 \log f(X_i;\lambda)}{\partial \lambda^2}\right) = \mathbb{E}\left(\frac{x}{\lambda^2}\right) = \frac{1}{\lambda^2}\mathbb{E}(x) = \frac{1}{\lambda}$$

سوال ۴

برای تست این مورد از آزمون Wald استفاده می‌کنم. فرض این آزمون با تعداد بالای نمونه تحقق پیدا می‌کند که به نظر می‌رسد ۱۰۰ نمونه کافی باشد. تخمین اولیه مطابق با MLE برای مقدار $\hat{p}$ برابر خواهد بود با 0.6. مطابق آزمون داریم:

$$\frac{\hat{p} - p_0}{\widehat{se}} \rightsquigarrow N(0,1)$$

مقدار $\widehat{se}$ هم برابر است با:

$$\widehat{se} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.6 * 0.4}{100}} \approx 0.049$$

پس نهایتا داریم:

$$W \approx \frac{0.6 - 0.5}{0.049} = 2.041$$

آزمون Wald در این سوال که یک طرفه (one-sided) هست با مقدار α زمانی رد می‌شود که داشته باشیم:

$$W > z_\alpha$$

الف) برای $\alpha = 0.05$ داریم:

$$z_\alpha = \Phi^{-1}(1-\alpha) = \Phi^{-1}(0.95) \approx 1.65$$

طبیعتا فرض H_0 رد می‌شود.

ب) برای $\alpha = 0.01$ داریم:

$$z_\alpha = \Phi^{-1}(1-\alpha) = \Phi^{-1}(0.99) \approx 2.33$$

فرض H_0 رد نمی‌شود.

ج) باید کمترین α را پیدا کرد که با آن H_0 رد شود پس:

$$W = z_\alpha \rightarrow 2.041 = z_{alpha} \rightarrow \Phi(2.041) = 1 - \alpha \rightarrow \alpha \approx 1 - 0.979 \rightarrow \alpha \approx 0.021$$

پس p-value در حدود 0.021 است.

سوال ۵

الف)

$$\mathbb{E}(\bar{X}_n^*|X_1, X_2, \dots, X_n) = \mathbb{E}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n}\sum \mathbb{E}(X_i) = \frac{1}{n}\sum \mathbb{E}(X) = \mathbb{E}(x)$$

ب)

$$\mathbb{V}(\bar{X}_n^*|X_1, X_2, \dots, X_n) = \mathbb{V}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum \mathbb{V}(X_i) = \frac{1}{n^2}\sum \mathbb{V}(X) = \frac{\mathbb{V}(X_i)}{n}$$

ج)

$$\mathbb{E}(\bar{X}_n^*) = \mathbb{E}\left(\frac{1}{n}\sum X_i^*\right) = \frac{1}{n}\sum \mathbb{E}(X_i^*) = \frac{1}{n}\sum \mathbb{E}(X) = \mathbb{E}(X)$$

د)

$$\mathbb{V}(\bar{X}_n^*) = \mathbb{E}\left({\bar{X}_n^*}^2\right) - \mathbb{E}(\bar{X}_n^*)^2 = \mathbb{E}\left(\left(\frac{1}{n}\sum X_i^*\right)^2\right) - \mathbb{E}(X)^2$$
$$= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n} \mathbb{E}\left(X_i^* X_j^*\right) - \mathbb{E}(X)^2$$

حال باید به قسمت $\mathbb{E}\left(X_i^* X_j^*\right)$ توجه کرد. اگر دو متغیر یکسان باشند، می‌توان این عبارت را به صورت $\mathbb{E}\left({X_i^*}^2\right)$ و سپس به $\mathbb{E}(X^2)$ تبدیل کرد. و اگر یکسان نباشند، این عبارت به $\mathbb{E}(X_i^*)\mathbb{E}\left(X_j^*\right)$ و سپس به $\mathbb{E}(X)^2$ تبدیل کرد. اگر $i = j$ باشد قطعا یک داده عینا تکرار شده است. ولی حتی اگر $i \neq j$ باشد باز محتمل است که یک داده عینا تکرار شده باشد، چراکه نمونه‌برداری با جایگزینی را در پیش گرفته‌ایم. با توجه به آنکه داده‌های اولیه برابر نیستند این احتمال برابر با $\frac{1}{n}$ خواهد بود. پس اگر i را ثابت فرض کنیم و jها را بررسی کنیم، یک جمله از n به علاوه $\frac{1}{n}$ از n-1 جمله دارای مقدار یکسانی است پس داریم:

$$\sum_{j=1}^{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i^*X_j^*\right) = \sum_{i=1}^{n}\left(\mathbb{E}(X^2) + \frac{n-1}{n}\mathbb{E}(X^2) + \frac{(n-1)^2}{n}\mathbb{E}(X)^2\right)$$
$$= (2n-1)\mathbb{E}(X^2) + (n-1)^2\mathbb{E}(X)^2$$

نهایتا داریم:

$$\mathbb{V}(\bar{X}_n^*) = \frac{1}{n^2}\left((2n-1)\mathbb{E}(X^2) + (n-1)^2\mathbb{E}(X)^2\right) - \mathbb{E}(X)^2$$
$$= \frac{2n-1}{n^2}\mathbb{E}(X^2) + \frac{-2n+1}{n^2}\mathbb{E}(X)^2 = \frac{2n-1}{n^2}\left(\mathbb{E}(X^2) - E(X)^2\right)$$
$$= \frac{2n-1}{n^2}\mathbb{V}(X)$$

قسمت پیاده‌سازی

سوال ۱

پاسخ