جلسه‌ی بیست و یکم آمار و احتمال مهندسی " دکتر شریفی زارچی - دانشگاه شریف "

یک اتفاقیست داریم که شامل $X_1 \dots X_n \sim N(\underbrace{\theta_1}_{\mu}, \underbrace{\theta_2}_{\sigma^2})$ داریم. ما می‌خواهیم که $MLE_{\hat{\theta}_1}, MLE_{\hat{\theta}_2}$ رو پیدا کنید.

$$\mathcal{L}(\theta_1, \theta_2 \mid x_1 \dots x_n) = f(x_1 \dots x_n; \theta_1, \theta_2)$$

$$= f(x_1; \theta_1, \theta_2)\, f(x_2; \theta_1, \theta_2) \dots f(x_n; \theta_1, \theta_2) = \prod_i \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}}$$

در واقع داریم روی یک جامعه بارها و بارها sample های مستقل از هم می‌گیریم و به دنبال پیدا کردن σ^2 در کل جامعه به شرط اون داده ها هستیم. اگر θ_1, θ_2 را بیابیم که حاصل ضرب تابع چگالی رو ماکزیمم کنه. احتمال اینکه اون دوتا پارامتر، همون پارامترهای جامعه باشن خیلی زیاد است. خودشون نیستن اما شانس بالایی دارن.

$$\underset{\theta}{\operatorname{argmax}}\ \mathcal{L}(\theta \mid X) = \underset{\theta}{\operatorname{argmax}}\ \ln(\mathcal{L}(\theta \mid X))$$

argmax: تابعی که θ ما رو بیان می‌کنه که $\mathcal{L}$ و $\ln \mathcal{L}$ هر دو ماکزیمم شده‌اند.

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}} = \underset{\theta_1, \theta_2}{\operatorname{argmax}} \left(\frac{-n}{2} (\ln 2\pi + \ln \theta_2) - \sum \left(\frac{x_i - \theta_1}{\theta_2}\right)^2 \right)$$

به دنبال max کردن این تابع با پارامترهای θ_1, θ_2 هستیم که حسب برحسب θ_1, θ_2 مشتق می‌گیریم و برابر صفر می‌گذاریم. و $\hat{\theta}_1, \hat{\theta}_2$ بدست میاد.

$$\Rightarrow \hat{\theta}_1 = \underbrace{\frac{\sum x_i}{n}}_{\text{میانگین}}, \quad \hat{\theta}_2 = \underbrace{\frac{\sum (x_i - \hat{\theta}_1)^2}{n}}_{\text{واریانس}}$$

هردوش با $n \to \infty$ به سمت Unbiased شدن میل می‌کنند که خب عملاً همون دوتا تخمین‌گر مادر و این زیباست.

خصوصیات (MLE)

به ازای هر وضعیتی برقرار نیستن اما خب شرایط ساده‌ای داره که تقریباً همیشه برقرارند.

اگر $n \to \infty$

- Unbiased $\qquad \lim_{n \to \infty} E(\hat{\theta}_{MLE}) = \theta$
- Consistant $\qquad \forall \varepsilon > 0; \ \lim_{n \to +\infty} \mathbb{P}(\hat{\theta}_{MLE} - \theta) > \varepsilon = 0$
- برای n های بزرگ

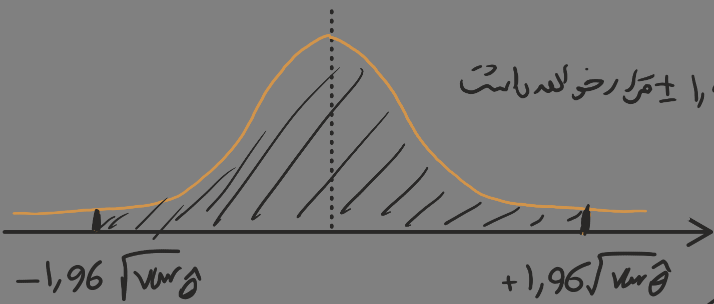
$$\frac{\hat{\theta}_{MLE} - \theta}{\sqrt{Var(\hat{\theta}_{MLE})}} \sim N(0,1)$$

جالبه که هر Estimator که MLE باشه همون میشه خیلی خفنه چون اسماً داره توزیعش رو هم همون میشه که خب در ادامه متوجه می‌شیم به چه دردی می‌خوره.

تا الان یک روشی (MLE) پیدا کردیم که به ما یک تخمینگر خوب میده.

خاصیت سوم MLE به چه دردی می‌خوره؟

اگر فرض کنیم که θ را همین بین و ما بیایم خاصیت سوم را اعمال کنیم. طبق خاصیت $N(0,1)$ خواهیم داشت



که ۹۵ درصد حالات بین اعداد $\pm 1,96\sqrt{var\,\hat{\theta}}$ قرار خواهد داشت

یعنی اگر همین یک θ بین ما با احتمال ۹۵ درصد مطمئنیم که اون θ زیاد دور از $\hat{\theta}$ ما نیست.

بعد از بدست آوردن $\hat{\theta}$. ما چقدر احتمال می‌دیم که $\theta = \hat{\theta}$، چون عضو $\mathbb{R}$ هستند صفر. چقدر احتمال داره که $\hat{\theta}$ به θ نزدیک باشه؟ نمی‌دونیم. حالا بهتره که $[\hat{\theta}_l, \hat{\theta}_h]$ بدیم که با احتمال ۹۵ درصد θ واقعی در این بازه هست.

تخمینگرهای بازه‌ای

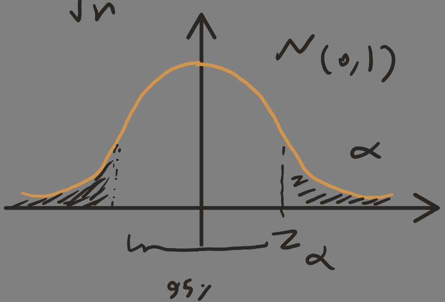
سوال ۱ فرض کنید $X_1 \dots X_n \sim N(\theta,1)$ برای $[\hat{\theta}_l, \hat{\theta}_h]$ رو پیدا کنید به شرطی که $\mathbb{P}(\theta \in [\hat{\theta}_l, \hat{\theta}_h]) \geq 95\%$

$$\bar{X} = \frac{X_1 + \dots + X_n}{n} \rightarrow var(\bar{X}) = n \times \frac{1}{n^2} = \frac{1}{n} \Rightarrow \bar{X} \sim N(\theta, \frac{1}{n})$$

جمع یک تعداد توزیع نرمال نرمال هست. با همین مر.

$$Z = \frac{\bar{X} - \theta}{\frac{1}{\sqrt{n}}} \sim N(0,1) \longrightarrow$$

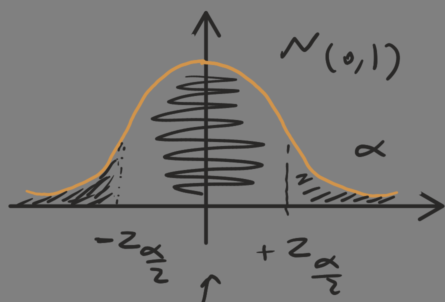
$\bar{X}$ رو نرمال سازی کردیم.



$$\mathbb{P}(Z > Z_\alpha) = \alpha$$

$$\phi^{-1}(1-\alpha) = Z_\alpha \rightarrow \alpha = 0.05$$

$$\mathbb{P}(\theta_l \leq Z \leq \theta_h) = \underbrace{1-\underbrace{\alpha}_{0.05}}_{95\%}$$



با احتمال ۹۵ درصد این وسط

$$\mathbb{P}\left(-Z_{\frac{\alpha}{2}} < \sqrt{n}(\bar{X}-\theta) < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Rightarrow \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}} + \bar{X} > \theta > \bar{X} - \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}}$$

$$\Rightarrow [\theta_l, \theta_h] = \left[\bar{X} - \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X} + \frac{Z_{\frac{\alpha}{2}}}{\sqrt{n}}\right], \quad Z_{\frac{\alpha}{2}} = ? \ ; Z_{0.025} = 1.96$$

$$\Rightarrow [\theta_l, \theta_h] = \left[\bar{X} - \frac{1.96}{\sqrt{n}}, \bar{X} + \frac{1.96}{\sqrt{n}}\right]$$

مفهوم این بازه چیه؟

ما براساس جمعیت n نفره یک میانگین ($\bar{X}$) پیدا کردیم. پس با احتمال ۹۵ درصد θ بین بازه‌ی $\bar{X} \pm \frac{1.96}{\sqrt{n}}$ هست. شاید هم به احتمال ۹۵ درصد بازه‌ی بدست آمده شامل θ باشد.

و مفهوم مهم‌تر: اگر $n \to \infty$ بازه‌ی ما تنگ‌تر و تنگ‌تر میشه و در نهایت به خود θ میرسیم. ما در نهایت تخفیف برای sample گرفتن نداریم. (اما با افزایش n وارد بازه‌ای میشیم که برای ما قابل قبوله.

سوال ۲

$$X_1 \dots X_n \sim \overset{?}{D}(\overset{?}{\theta}, \sigma^2)$$

$$[\hat{\theta}_l, \hat{\theta}_h] \ \Big| \ \mathbb{P}\left(\theta \in [\bar{X} - ?, \bar{X} + ?]\right) \geq 1-\underbrace{\alpha}_{\text{confidence}}$$

میزان احتمال

$$Z = \frac{\bar{X}-\theta}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

قضیه حد مرکزی. وقتی n بزرگ باشد آنگاه بازه‌ی اطمینان $1-\alpha$

$$\mathbb{P}\left(-Z_{\frac{\alpha}{2}} \leq Z \leq +Z_{\frac{\alpha}{2}}\right) \leq 1-\alpha \quad \Rightarrow \left[\bar{X} - \frac{Z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}, \bar{X} + \frac{Z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right]$$

مثال فرض کنید یک توزیع داریم با 30 = σ^2 برای یک بازه‌ی اطمینان ٪99 پیدا کنید.

$1-\alpha = .99 \rightarrow \alpha = 0.01 \rightarrow \frac{\alpha}{2} = 0.005$ $\quad Z_{\frac{\alpha}{2}} = \phi^{-1}(1-0.005) = 2.57$

$$\left[\bar{X} - \frac{2.57 \times \sqrt{30}}{\sqrt{n}} , \bar{X} + \frac{2.57 \times \sqrt{30}}{\sqrt{n}}\right]$$

خب خود σ^2 رو از کجا بیاریم؟

سؤال 3 $X_i \sim \overset{?}{D}(\overset{?}{\theta}, \overset{?}{\sigma^2})$

○ یک حد بالا روی σ : $P(\sigma \leq \sigma_{max}) = 1$: کلا روشی برای این هست مثل همین روش بازه‌ی اطمینان و...

○ تخمین σ^2 . $\underbrace{S^2 = \frac{1}{n-1} \Sigma (X_i - \bar{X})^2}$ ← اگر n بالا بود کار جمع لست است چون Unbiased هست

n بار اینو بدست میاریم و Max اون n بار رو گزارش می‌کنیم که $P(\sigma^2 \leq \sigma^2_{max})$ برابر 1 می‌شود

اگر n بار σ^2 را تخمین بزنیم با احتمال $\frac{1}{2^n}$, σ^2 واقعی از σ^2_{max} بزرگتر خواهد بود که دیگه خوبه خیلی