Whiteboard 7993728-4386-1506

فراسینه های مقدماتی : جلسه 26 و 27 خرداد 1400

---

* مطالب این جلسه :

- تخمینگرهای بیز (Bayes Estimators)
- روش های ارزیابی تخمینگرها
- خطای میانگین مربعات (Mean Square Error)

- تخمینگرهای بیز :

دو نگاه در آمار:
- نگاه کلاسیک ← θ پارامتر مدل و نامعلوم
- نگاه بیزین ← θ پارامتر مدل و نامعلوم ولی توزیعی یک متغیر تصادفی است.

↳ در نگاه بیز استفاده بیشین در مورد θ داریم که این مسئله را با فرض وجود توزیع برای θ نشان می دهیم ← توزیع بیشین prior dist.

(انتخاب ذهنی (به صورت subjective) (نقطه نظر نگاه بیز)

Whiteboard 7993728-4386-1506

فرض: یک توزیع پیشین $\pi(\theta)$ داریم که دانش اولیه درمورد پارامتر θ را بیان می‌کند.

پس از مشاهده مشاهدات ← توزیع پسین (Posterior dist.):

$$\pi(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)\cdot\pi(\theta)}{m(\underline{x})}$$

$f(\underline{x}|\theta)\cdot\pi(\theta) \rightarrow f(\underline{x},\theta)$

$m(\underline{x}) \rightarrow$ توزیع حاشیه‌ای $\underline{x}$

قاعده مدل‌های مورد بررسی

$$m(\underline{x}) = \int f(\underline{x},\theta)\,d\theta = \int f(\underline{x}|\theta)\cdot\pi(\theta)\,d\theta$$

با داشتن توزیع پسین θ ← درمورد θ نتیجه‌گیری انجام دهیم.

به عنوان مثال مثلا $\mathbb{E}[\theta|\underline{x}]$ به عنوان تخمینگر مقداری برای θ در نظر می‌گیریم.

مثال: فرض کنید $Y \sim Binomial(p,n)$ که n مقدار مشخص و p پارامتر نامعلوم است.

می‌خواهیم یک تخمینگر بیزی برای p بدست بیاوریم.

Whiteboard 7993728-4386-1506

فرض می کنیم که p توزیع بتا دارد.

$$\pi(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}, \quad \alpha,\beta>0, \quad 0\le p\le 1$$

$$\Rightarrow \mathbb{E}[p] = \frac{\alpha}{\alpha+\beta}$$

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x}\, dx$$

$$\Gamma(t+1) = t\Gamma(t), \quad \Gamma(1)=1$$

$$\Rightarrow \Gamma(k+1) = k!$$

اگر عدد صحیح مثبت

حال اگر متغیر y توزیع p به شکل زیر توابعی بود:

$$\pi(p|y) = \frac{f(y,p)}{f(y)}$$

$$f(y,p) = f(y|p)\cdot\pi(p) = \left[\binom{n}{y} p^y (1-p)^{n-y}\right]\cdot\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}$$

Whiteboard 7993728-4386-1506

$$= \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1} (1-p)^{n-y+\beta-1}$$

$$\Rightarrow f(y) = \int_0^1 f(y,p)\cdot dp = \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} \underbrace{\int_0^1 p^{y+\alpha-1} (1-p)^{n-y+\beta-1} dp}$$

$$= \frac{\Gamma(y+\alpha)\,\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)}$$

$$\Rightarrow \pi(p|y) = \frac{f(y,p)}{f(y)} = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\,\Gamma(n-y+\beta)} p^{y+\alpha-1} (1-p)^{n-y+\beta-1}$$

توزیع پسین هم یک توزیع

خانواده

بتا سازگار یا رابطه‌ای دیگر ← توزیع بتا (conjugate family) برای توزیع دوجمله‌ای.

با درنظرگرفتن توزیع پیشین به عنوان مثال امید ریاضی را به عنوان نمونه مقدار p در نظر

می‌گیریم:

Whiteboard 7993728-4386-1506

$$\hat{P}_B = \mathbb{E}[P \mid y] = \frac{y+\alpha}{n+\alpha+\beta}$$

$$\hat{P}_{ML} = \frac{y}{n}$$ (از جلسه ی قبل):

$$\hat{P}_B = \frac{n}{n+\alpha+\beta}\underbrace{\left(\frac{y}{n}\right)}_{\hat{P}_{ML}} + \frac{\alpha+\beta}{n+\alpha+\beta}\underbrace{\left(\frac{\alpha}{\alpha+\beta}\right)}_{\mathbb{E}[P]}$$

$\mathbb{E}[P] \sim \lambda(P)$

مثال: فرض کنید $X \sim N(\theta, \sigma^2)$ و $\theta \sim N(\mu, \tau^2)$

σ^2: معلوم ← نامعلوم

μ, τ^2: معلوم

$$\lambda(\theta) = \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(\theta-\mu)^2}{2\tau^2}\right)$$

Whiteboard 7993728-4386-1506

$$\pi(\theta|x) = \frac{f(x|\theta)\cdot\pi(\theta)}{f(x)}$$: و تولن ستان دار توزیع پسین

هر گاوسی ی کنور با پارامترهای:

$$\hat{P}_B \rightarrow \begin{cases} E[\theta|x] = \frac{\tau^2}{\tau^2+\sigma^2}x + \frac{\sigma^2}{\tau^2+\sigma^2}\mu \\ Var[\theta|x] = \frac{\sigma^2\tau^2}{\sigma^2+\tau^2} \end{cases}$$

$x \rightarrow \hat{P}_{ML}$

$\mu \rightarrow$ پیشینی متغلد برحسب دانش اولیه

---

* روش هایی برای ارزیابی تخمینگر ها:

• خطای میانگین مربع (Mean Square Error):

تعریف: خطای MSE یک تخمینگر W برای پارامتر θ و تابعی از مقدار θ که به صورت زیر تعریف می شود:

$$e_{MS}(\theta) = E[(W-\theta)^2]$$

معیارهای مشابه برای خطا : $E[f(|W-\theta|)]$

Whiteboard 7993728-4386-1506

معیار MSE خواص دو خاصیت خوب دارد :

① بسیاری از مسایل به صورت تحلیلی قابل بررسی هستند.

② قضیه زیر هم برای آن برقرار است :

$$e_{MS}(\theta) = \mathbb{E}[(W-\theta)^2] = Var[W] + (\mathbb{E}[W]-\theta)^2$$
$$= Var[W] + (Bias[W])^2$$

اگر خطای MSE یک تخمینگر کم باشد ← هم $Var[W]$ کم و هم $Bias[W]$ کم هستند.

تخمینگر unbiased یا بدون بایاس : $\mathbb{E}[W] = \theta$

مثال: $X_1, \dots, X_n$ متغیرهای iid گاوسی با توزیع $\mathcal{N}(\mu, \sigma^2)$.

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} \quad \Rightarrow \mathbb{E}[\bar{X}] = \mu$$

Whiteboard 7993728-4386-1506

$$S^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2 \Rightarrow \mathbb{E}[S^2] = \sigma^2$$

می‌توان بررسی کرد که در حالت کلی هر دو این تخمگر unbiased هستند.

با فرض گاوسی بودن برای خطای MSE این تخمگر‌ها داریم:

$$\mathbb{E}[(\bar{X} - \mu)^2] = Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$\mathbb{E}[(S^2 - \sigma^2)^2] = Var[S^2] = \frac{2\sigma^4}{n-1}$$ (از گاوسی بودن استفاده کردیم)

از جلسات قبل برای تخمگر ML برای σ^2 داشتیم:

$$\hat{\sigma}^2_{ML} = \frac{n-1}{n} S^2 = \frac{1}{n}\sum (X_i - \bar{X})^2 \Rightarrow \mathbb{E}[\hat{\sigma}^2_{ML}] = \frac{n-1}{n}\sigma^2$$ biased

$$\Rightarrow Var[\hat{\sigma}^2_{ML}] = \left(\frac{n-1}{n}\right)^2 Var[S^2] = \frac{2(n-1)\sigma^4}{n^2}$$

Whiteboard 7993728-4386-1506

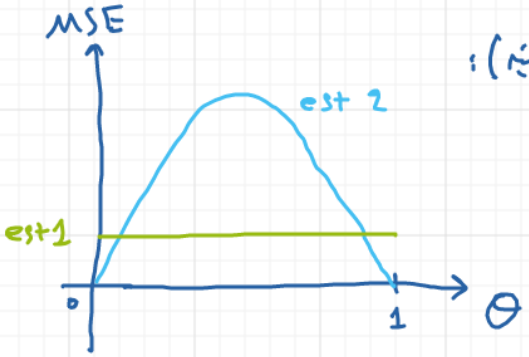
MSE

$$\Rightarrow \mathbb{E}\left[\left(\hat{\sigma}^2_{ML} - \sigma^2\right)^2\right] = \mathrm{Var}\left[\hat{\sigma}^2_{ML}\right] + \mathrm{Bias}\left[\hat{\sigma}^2_{ML}\right] = \frac{2n-1}{n^2}\sigma^4$$

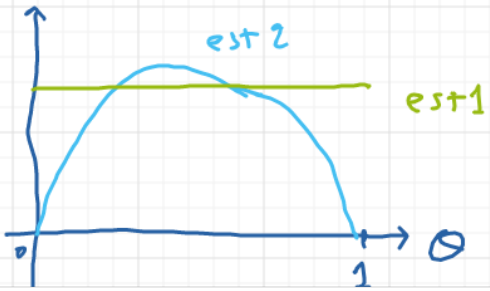
$$\Rightarrow \mathbb{E}\left[\left(\hat{\sigma}^2_{ML} - \sigma^2\right)^2\right] = \frac{2n-1}{n^2}\sigma^4 < \frac{2\sigma^4}{n-1} = \mathbb{E}\left[\left(S^2 - \sigma^2\right)^2\right]$$

↑ biased

↑ unbiased



مثال کلی (عالی اصل بکار بردن مرا میفهمید):



Whiteboard 7993728-4386-1506

اقدام تنبیه 1 را انتخاب می‌کنیم.
مگر اینکه اطلاعات پیشین داشته باشیم
که θ خیلی نزدیک به 0 یا 1 است.

اقدام تنبیه 2 را انتخاب می‌کنیم
مگر اینکه میانه θ خیلی نزدیک $\frac{1}{2}$ است.

⇐ برای یک مسئله یک تنبیه بهینه مطلق برای هر بردار θ وجود نداشته باشد.
ممکن است