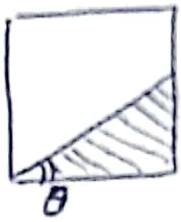
۱) ابتدا تابع CDF، یعنی $F_\Theta(\theta)$ را بدست می‌آوریم.

بدین منظور، ابتدا فرض کنید که داشته باشیم $\theta \le \frac{\pi}{4}$، در این صورت داریم:

$$F_\Theta(\theta) = \mathbb{P}\left[\angle(X,Y) \le \theta\right] =$$



مساحت ناحیه هاشور خورده.

$$\Rightarrow \quad F_\Theta(\theta) = \frac{\tan\theta \times 1}{2} = \frac{1}{2}\tan\theta$$

$\theta \le \frac{\pi}{4}$

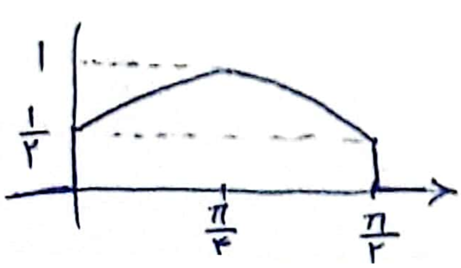
$$f_\Theta(\theta) = \frac{dF_\Theta(\theta)}{d\theta} = \frac{1}{2}\left(1 + \tan^2\theta\right) = \boxed{\frac{1}{2\cos^2\theta}}$$

از طرفی، مسئله نسبت به $\frac{\pi}{4}$ تقارن دارد. در نتیجه، داریم:

$$f_\Theta(\theta) = f_\Theta\left(\frac{\pi}{2} - \theta\right)$$

$0 \le \theta \le \frac{\pi}{4}$

$$\Rightarrow \quad \boxed{f_\Theta(\theta) = \begin{cases} \frac{1}{2\cos^2\theta} & 0 \le \theta \le \frac{\pi}{4} \\ \frac{1}{2\sin^2\theta} & \frac{\pi}{4} \le \theta \le \frac{\pi}{2} \end{cases}}$$



۲) الف)

$$P(X_i = x) = p^{x}(1-p)^{1-x}$$

$$\Rightarrow P(\text{Data} \mid p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= p^{\left(\sum_i x_i\right)}(1-p)^{n-\left(\sum_i x_i\right)} = p^{m}(1-p)^{n-m}$$

$$\Rightarrow LLH(p) = m \log p + (n-m)\log(1-p)$$

باید بیشینه LH یا LLH را پیدا کنیم.

$$\frac{d\,LLH(p)}{dp} = \frac{m}{p} - \frac{n-m}{1-p} = 0$$

$$\Rightarrow m(1-p) = p(n-m)$$

$$\boxed{\hat{p}_{MLE} = \frac{m}{n}}$$

ب)

$$\hat{p}_{MAP} = \underset{0 \le p \le 1}{\operatorname{argmax}}\ \mathbb{P}(\text{Data} \mid p)\,\mathbb{P}(p)$$

یا

$$f(\text{Data} \mid p)\, f(p)$$

یا

$$\mathbb{P}(\text{Data} \mid p)\, f(p)$$

چون داده‌ها گسسته هستند و p پیوسته

$$\Rightarrow f(p) = \frac{1}{2}\,\text{Unif}(0,\alpha) + \frac{1}{2}\,\delta(p-\beta)$$

$$= \frac{1}{2\alpha}\,\mathbb{1}(0 \le p \le \alpha) + \frac{1}{2}\,\delta(p-\beta)$$

$$\Rightarrow \hat{p}_{MAP} = \underset{p}{\arg\max}\ \mathbb{P}(Data|p)\left[\frac{1}{2\alpha}\mathbb{1}(0\le p\le\alpha) + \frac{\delta(p-\beta)}{2}\right]$$

روی β جهش بی‌نهایت است.

$$\Rightarrow \boxed{\hat{p}_{MAP} = \beta}$$

از طرفی $0<\beta<1$، پس

$$P(Data|\beta) \neq 0$$

ج) تخمین‌گر MAP در این مسئله مستقل از داده‌ها است. که خوب نیست.
می‌توان به جای MAP از تخمین‌گر «امید ریاضی پسین» استفاده کرد.

$$\Rightarrow \hat{p} \triangleq \int_0^1 p\, f(p|Data)\,dp$$

$$f(p|Data) = \frac{P(Data|p)\,f(p)}{\mathbb{P}(Data)}$$

$$f(p|Data) \xrightarrow{\text{وقتی Data نداریم}} = f(p) \quad \text{prior}$$

$$\xrightarrow[\text{وقتی } n\to\infty]{} \overset{a.s.}{=} \delta\left(p-\frac{m}{n}\right) \quad \text{likelihood.}$$

۳) الف) تخمینگر پیشنهادی کوواریانس نمونه‌ای است:

$$\boxed{\hat{\sigma} \triangleq \frac{1}{n}\sum_{i=1}^{n} X_i Y_i}$$

به خاطر قانون اعداد بزرگ، به $\mathbb{E}(XY)$ میل می‌کند، پس سازگار است.

$$\mathbb{E}\,\hat{\sigma} = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i Y_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sigma_{XY} = \sigma_{XY}$$

پس بایاس ندارد

دقت کنید که چون $\mu_X = \mu_Y = 0$، پس داریم

$$\mathbb{E}(X_i Y_i) =$$

$$\mathbb{E}\big((X_i - \mu_X)(Y_i - \mu_Y)\big) = \sigma_{XY}$$

ب) در این حالت، میانگین‌ها دیگر صفر نیستند.

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

$$\hat{\mu}_Y \triangleq \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$\hat{\sigma} \triangleq \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu}_X)(Y_i - \hat{\mu}_Y)$$

به خاطر این عبارات، قانون اعداد بزرگ قابل اعمال نیست.

$$\mathbb{E}\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[(X_i-\hat{\mu}_X)(Y_i-\hat{\mu}_Y)\right]$$

$$X_iY_i - X_i\hat{\mu}_Y - Y_i\hat{\mu}_X + \hat{\mu}_X\hat{\mu}_Y$$

I) $\mathbb{E}(X_iY_i) = \mathbb{E}\big((X_i-\mu_X+\mu_X)(Y_i-\mu_Y+\mu_Y)\big) =$

$$\sigma_{XY} + 0 + 0 + \mu_X\mu_Y = \sigma_{XY} + \mu_X\mu_Y$$

II) $\mathbb{E}(X_i\hat{\mu}_Y) = \mathbb{E}\left(\frac{1}{n}X_iY_i\right) + \frac{1}{n}\sum_{\substack{j=1 \\ j\neq i}}^{n}\mathbb{E}(X_iY_j)$

$$= \frac{1}{n}\sigma_{XY} + \frac{\mu_X\mu_Y}{n} + \frac{n-1}{n}\mu_X\mu_Y =$$

$$\frac{1}{n}\sigma_{XY} + \mu_X\mu_Y$$

III) $\mathbb{E}(Y_i\hat{\mu}_X) = \frac{1}{n}\sigma_{XY} + \mu_X\mu_Y \longrightarrow$ II) به طریق مشابه

IV) $\mathbb{E}(\hat{\mu}_X\hat{\mu}_Y) = \frac{1}{n^2}\sum_{i,j=1}^{n}\mathbb{E}(X_iY_j) = \frac{n}{n^2}\sigma_{XY} + \frac{n^2}{n^2}\mu_X\mu_Y$

$$= \frac{1}{n}\sigma_{XY} + \mu_X\mu_Y$$

I) + II) + III) + IV)

$$\Rightarrow \mathbb{E}\hat{\sigma} = \sigma_{XY} + \frac{1}{n}\sigma_{XY} - \frac{2}{n}\sigma_{XY} = \boxed{\sigma_{XY}\left(1-\frac{1}{n}\right)}$$

پس بایاس دارد. ولی به سمت صفر می رود.

۴)

$$X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_n$$

$$x_i \sim Unif(0,1) \Rightarrow \mathbb{P}\left[\begin{matrix} \text{حداقل } 1\ x_i \\ \text{بین } (1-\epsilon, 1] \end{matrix}\right] =$$

$$1 - \mathbb{P}\left[\text{همگی بین } [0, 1-\epsilon]\right] = 1 - (1-\epsilon)^n$$

الف) و ب) حل شدند.

ج)

$$1-(1-\epsilon)^n \simeq 1 - e^{-n\epsilon}$$

$$\epsilon \triangleq \frac{1}{\sqrt{n}}$$

$$\Rightarrow \mathbb{P}\left(\begin{matrix} \text{حداقل یک } x_i \text{ در فاصله} \\ \frac{1}{\sqrt{n}} \text{ از مرز} \end{matrix}\right) \cong 1 - e^{-n \times \frac{1}{\sqrt{n}}} = 1 - e^{-\sqrt{n}}$$

د)

$$x_i \sim \mathcal{N}(0,1) \Rightarrow \mathbb{E} X_i^2 = \mathbb{E}\left[(X_i - \mu)^2\right] = Var(X_i) = 1$$

چون $\mu = 0$

ه)

$$Var(X_i^2) = \mathbb{E}\left[(X_i^2 - \mathbb{E}X_i^2)^2\right] = \mathbb{E}\left((X_i^2 - 1)^2\right) =$$

$$\mathbb{E}\left(X_i^4 - 2X_i^2 + 1\right) = 3 - 2 + 1 = 2$$

و)

$$R^2 = \underbrace{X_1^2 + \cdots + X_n^2}_{} \Rightarrow \begin{matrix} \mathbb{E}R^2 = n \\ Var(R^2) = 2n \end{matrix}$$

جمع یک سری متغیر i.i.d.

میانگین و واریانس جمع می‌شوند

$$\Rightarrow \mathbb{P}\left[|R^2 - n| > k\sqrt{2n}\right] \le \frac{1}{k^2}$$

Chebyshev's Inq.

$$k = n^{1/4}$$

$$\Rightarrow \quad \mathbb{P}\left[|R^2 - n| > \sqrt{2}\, n^{3/4} \right] \leq \frac{1}{\sqrt{n}}$$

یا

$$\boxed{\mathbb{P}\left[n - \sqrt{2}\, n^{3/4} \leq R^2 \leq n + \sqrt{2}\, n^{3/4} \right] \geq 1 - \frac{1}{\sqrt{n}}}$$

اثبات شد!