بسم الله الرحمن الرحیم

جلسه ۱۴۰۰/۸/۱۴

* Data modeling

$\{D_i\}, i=1, \ldots, N$ Data observation

اصول موضوعه / بدیهیات شناختی: model's free parameter ← $\{\theta_j\}, j=1, \ldots, M$

Merit function

$N > M$

☆ Bayesian approach

• کمیت های آزاد مدل $\{\theta\}$ مجهول و ما تنها به صورت احتمالی مقدار آنها را می توانیم تعیین کنیم. بازه ای که مقدار بهینه پارامتر در آن وجود دارد تعیین می شود.

*• سازوکار خوش تعریفی برای استفاده از اطلاعات اولیه (Prior information) به منظور تعیین مقادیر بهینه پارامترهای آزاد وجود دارد.

به بیان دیگر اطلاعات در یک مرحله از انجام اندازه گیری به منظور تعیین مقدار بهینه پارامتر به عنوان اطلاعات اولیه هنگامی که در یک مرحله جدیدی اندازه گیری به منظور تعیین مقدار پارامتر انجام شود قابل استفاده است.

Marginalization (Integrate out)

• سازوکار خوش تعریفی برای بی اثرسازی اثر پارامترهای زائد وجود دارد

وجود سازوکار حذف یا بی اثرسازی یک یا چند پارامتر آزاد مدل

Marginalization

مفهوم ترکیب نسبت به کمیت پارامترهای زائد

☆ Frequentist approach

• کمیت آزاد مدل $\{\theta\}$ مجهول ولی دارای مقدار منحصر و ثابت هستند

• هیچ گونه سازوکار خوش تعریفی برای استفاده از اطلاعات مربوط به مدل (نظری) برای تعیین مقدار بهینه پارامترها وجود ندارد.

• هیچ نوع اطلاعاتی از مدل مورد استفاده نیست.

{Prior information} بلااستفاده است.

• سازوکار خوش تعریفی برای رهایی از پارامترهای اضافی

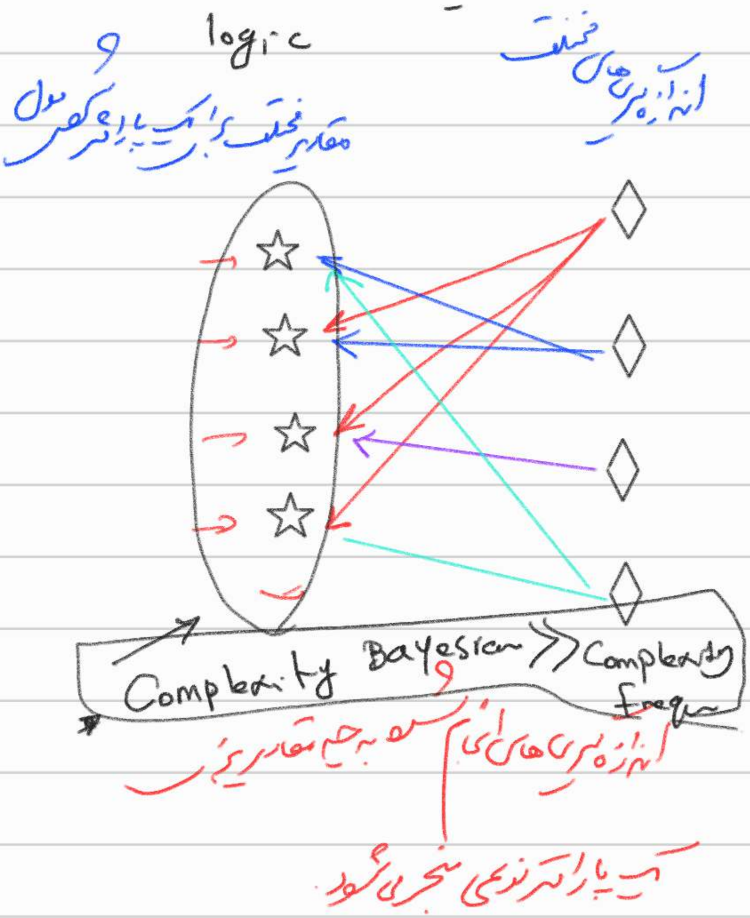
Nuisance Parameters

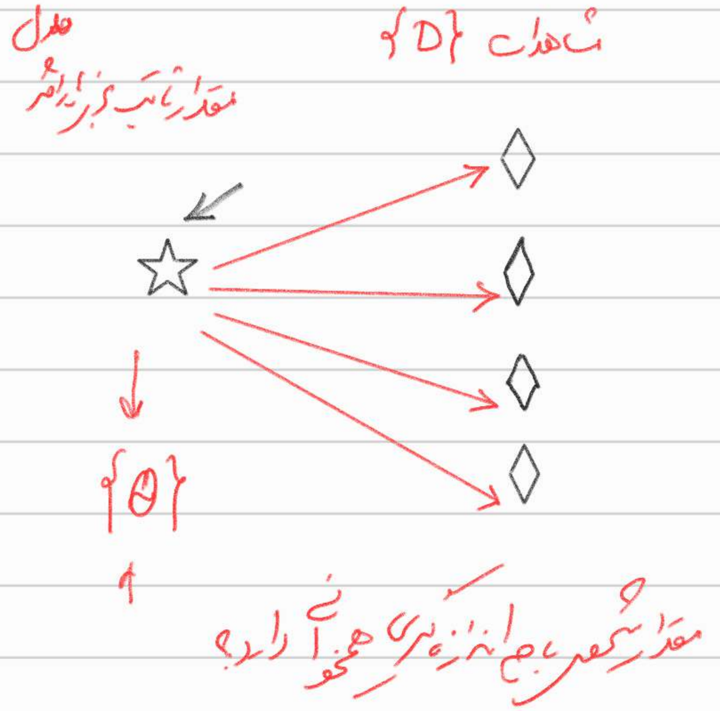
وجود ندارد (وجودشان با اندازه گیری ها بی تاثیر است)

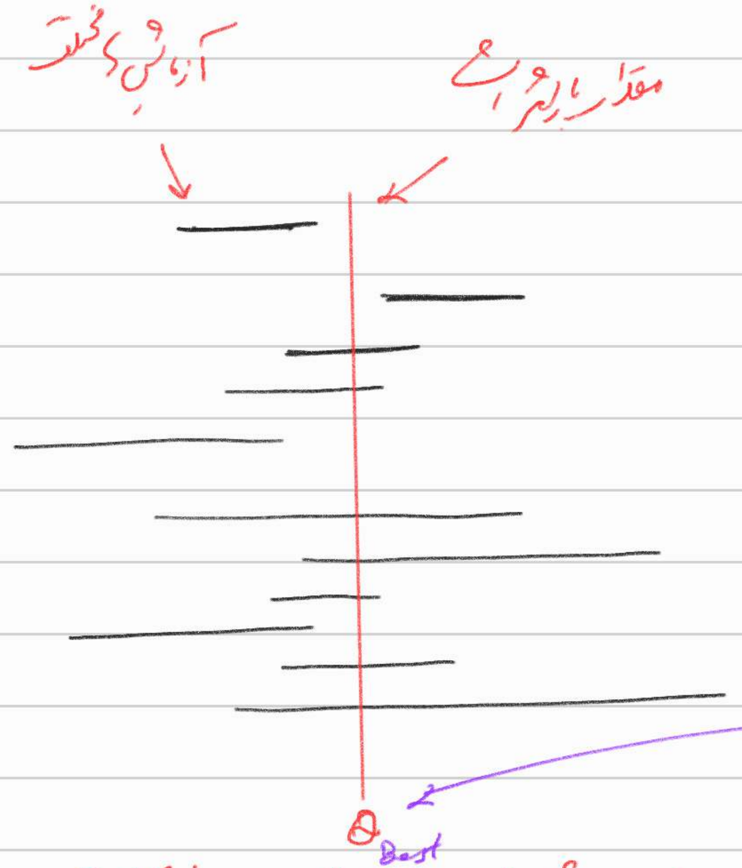
در نتیجه گیری ← با اندازه گیری های موجود امکان تعیین مقادیر بهینه آنها وجود ندارد)

○ در رویکرد بیزین ما با روش استقرایی موضوع را پی می‌گیریم Inductive logic

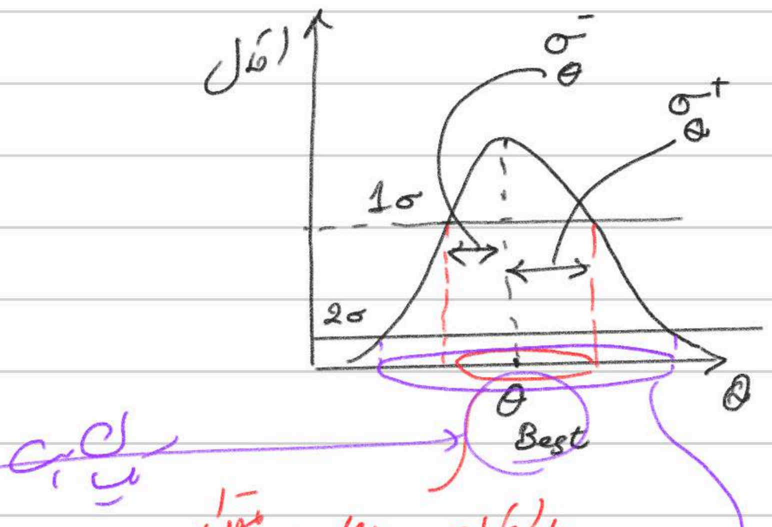
○ در رویکرد Frequentist منطق قیاسی درنظر گرفته می‌شود.







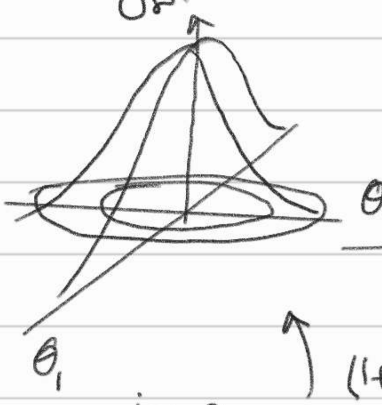
۹۰ درصد شاهدات به مقدار مشخص θ منجر شود.



با احتمال ۶۸ درصد مقدار وقوع پارامتر در این ناحیه است.

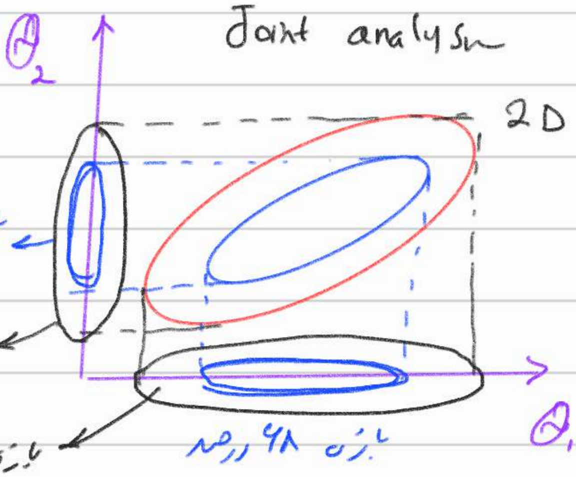
با احتمال ۹۵ درصد مقدار وقوع پارامتر در این ناحیه است.

اول



(1+2)

$j=1,2$



M → (1+M) - Dimension

M-Dimension

$$\{\theta\} = \{\theta_{Best}\} \pm \{\sigma_\theta\}$$

$$= \{\theta_{Best}\}^{+\{\sigma_\theta^+\}}_{-\{\sigma_\theta^-\}}$$

Degree of freedom — درجه آزادی

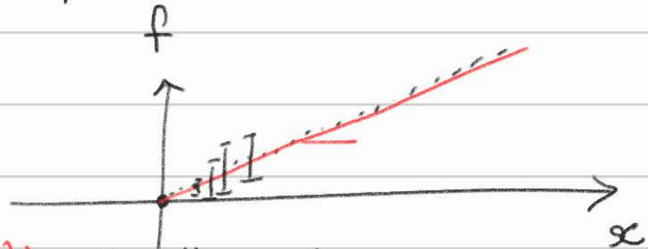
$$\nu \equiv DoF \equiv N - M$$

همه این پارامترهای آزاد مدل هستند — $\{\theta_j\}, \; j=1, \dots M$

اول

دوم

$$f = mx + b$$



$m \gg b$

$$\begin{cases} m = \mathcal{O}(10^{2}) + \mathcal{O}(10^{-1}) \\ b \sim \mathcal{O}(10^{-4}) \mp \mathcal{O}(10^{-2}) \end{cases}$$

→ معیار سرانگشتی برای پارامتر اضافی

عدد دقیق

b خوب مقید نمی‌شود.

Mathematical Description of Bayesian method

① $\{D\}_{i=1,N}$: observed Data (Experiment)
آزمایش

$\{\Theta\}_{j=1,M}$: Free Parameter → $\{\theta\}=$ ✓
از مدل آمده

Maximization of Posterior function
تابع پسینی

$$\boxed{P(\{\Theta\} \mid \{D\})}$$

احتمال بدست آمدن مقدار $\{\theta\}$ اگر $\{D\}$ بدست آمده باشند.

Conditional Probability function
تابع احتمال شرطی

$$P_{max}(\{\Theta\}=\{\Theta_{Best}\} \mid \{D\})$$

↑

Best fit Values for free Parameter are associated with maximum Value of Posterior

② Bayse theorem
قضیه بیز

$$P(\{\theta\} \mid \{D\}) = \frac{P(\{D\} \mid \{\theta\})\, P(\{\theta\})}{\int d\{\theta\}\, P(\{\theta\}, \{D\})}$$

$$= \frac{p(\{D\}|\{\theta\})\, p(\{\theta\})}{\int d\{\theta\}\, p(\{D\}|\{\theta\})\, p(\{\theta\})}$$

likelihood

احتمال اینکه $\{\Theta\}$ به آنچه به شرط اینکه $\{D\}$ رخ داده شوند $=$ $\dfrac{\text{احتمال پیش آمدن داده ها به شرط مدل} \times \text{احتمال مدل}}{\text{Normalization factor}}$

Prior

Posterior

اطلاعات اولیه در خصوص مدل است

$$P(A|B) = \frac{P(A,B)}{P(B)} =$$

چنین محتمل هستند

$$P(\{\theta\}|\{D\}) \simeq \mathcal{L}(\{D\}|\{\theta\})\; p(\{\theta\})$$

Posterior — Likelihood — Prior

مثال عددی

$\{\theta\}: \{a, b\}$

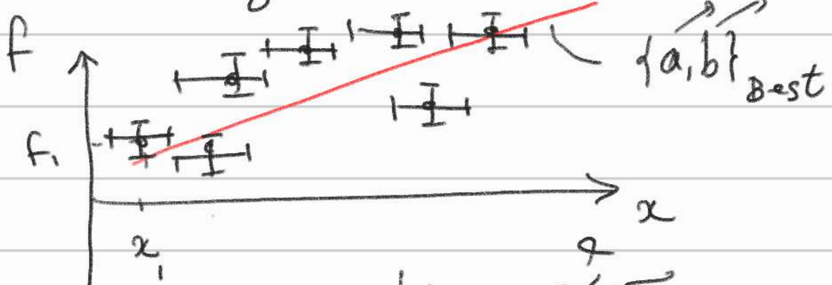
$f(x) = ax + b$ $\quad j = 1, 2$

$M = 2$

$\{D\}: \{(x_1, f_1); (x_2, f_2) \dots\}$, $i = 1, \dots, N$

x	f
x_1	f_1
x_2	f_2
$\vdots$	
x_N	f_N

$f_{Theory}(x; \{\theta\}) = (x, \{a, b\})$



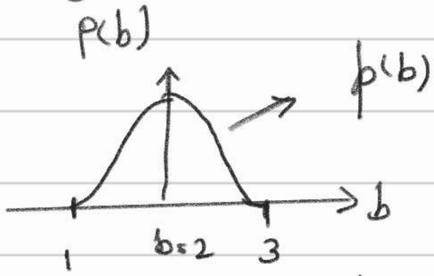
با توجه به نوع آزمایش $b > 0$ (انتظار: اطلاعات اولیه، علم لدنی)

$p(b < 0) = 0$

* انتظار می رود که $1 < b < 3$

$$p(b) = \begin{cases} 0 & b \leq 1 \\ cts & 1 < b < 3 \\ 0 & b \geq 3 \end{cases}$$

* انتظار می رود $1 < b < 3$ و حول $b = 2$ با احتمال بیشتری قرار دارد.



کلیه اطلاعات اولیه مربوط به پارامترهای مدل که به طور کلی از طریق اندازه گیری به دست نیامده است با توجه به شکل اطلاعات پیشینی (Prior information) قابل استفاده در مدل سازی Bayesian است.

اطلاعات اولیه می توانند از اندازه گیری به دست آید: مثلاً ما امروز در آزمایشگاه ابن سینا

V بر حسب جریان I اندازه گیری کنیم،

$a \leftrightarrow R$ ، $I \leftrightarrow x$ ، $V \leftrightarrow F$

I	V
I_1	V_1
⋮	
I_{100}	V_{100}

$\longrightarrow R = \checkmark$

Prior information

(فرآیندها وضعیت قبل (تقریبی

I	V
I_1'	V_1'
⋮	
I_{500}'	V_{500}'

$\longrightarrow R' = \checkmark$

I	V
I_1	
I_{100}	V_{100}
I'_1	V'_1
I'_{500}	V'_{500}

تا 1,600 → R s✓

این کار معادل اینه که

improvement in observation

فضای احتمال پارامترها در داده‌ها

Likelihood

اغلب همه سعی می‌کنند در سازوکارهای بیشتر از این دیدگاه رهیافت درست‌نمایی را بیشینه کنند.

$$\mathcal{L}(\{D\}|\{\theta\}) \rightarrow \text{Maximization}$$

اغلب استفاده می‌شود

$$p(\{\theta\}|\{D\}) = \frac{\mathcal{L}(\{D\}|\{\theta\})\, \underbrace{p(\{\theta\})}_{1}}{\int d\{\theta\}\, \mathcal{L}(\{D\}|\{\theta\})\, \underbrace{p(\{\theta\})}_{1}}$$

[illegible] →

اغلب $p(\{\theta\}) = 1$ ← یعنی هیچ ارجحیت خاصی بر روی مقادیر پارامترها قبل از انجام عمل آزمایش کردن تابع درست‌نمایی در نظر نگرفته‌ایم.

$$\text{Posterior} \simeq \text{Likelihood}$$

* الان وقت آن است که شکل ریاضیاتی $\mathcal{L}$ و Posterior را بسازیم

سپس با یک مثال در خصوص همین Posterior Analysis خواهیم دانست و سپس به سوال بالا پاسخ می‌دهیم.