

* البته می‌توانیم به β نرمال چند متغیره هم بدیم ولی در دسترس داریم.

$$\beta \sim N \left(\begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} 1 & & & & \\ & \dots & & & \\ & & 1 & & \\ & & & \dots & \\ & & & & 1 \end{bmatrix} \right)$$

$$\text{Beta} \sim \text{dmnorm}(\text{muo}[1:5], \text{Lambda}[1:5, 1:5])$$

5

```
for (r in 1:5) { muo[r] <- 0 }
```

#data ↓
داده

```
list(N=48, structure(
```

10

از توی مثال‌ها ببین!

* خلاصه بسیار «معیدهای انتخاب مدل» ۱۳، ۹۴ - شماره

اساس انتخاب مدل بر اساس χ^2 حرف:

15

<p>این روش‌ها Free از توزیع هستند (توزیع: λ ها بلیک است)</p> <p>فقط تعداد Covariate ها عوض می‌شوند. در واقع Variable Selection است</p>	}	<p>① نیدوی بزرگش (مثل R^2)</p> <p>② کمترین خطای پیش بینی (مثل CV, CP)</p>
---	---	---

<p>در این روش‌ها توزیع‌ها هم می‌توانند فرق کنند.</p>	}	<p>③ بر اساس likelihood</p> <p>④ بر اساس posterior ²⁰</p>
--	---	--

در این‌ها علاوه بر این که اطلاعات مربوط به تعداد متغیرها وجود دارد، اطلاعات در شانس و مدل هم وجود دارد.

25

۱- معیار AIC : این معیار من لوید اگر $f(y)$ توزیع واقعی داده ها باشد و $g(y)$

مدل پیشنهادی ، آن $g(y)$ و این که فاصله کولب - لیب (Kullback-Leiber) را \min کند بهترین مدل است .

$$KL(f, g) = \int \log \frac{f(y)}{g(y)} f(y) dy$$

بر این اساس و پس از انجام محاسبات ، آن مدلی که KL را \min کند همان مدلی

است که عبارت زیر (یعنی معیار AIC) را \min کند

$$AIC = -2 \log g(y, \hat{\theta}(y)) + 2r$$

MLE of θ بعد بردار پارامتر θ

اطلاعات درستی مدل + حجم جریمه
(measure of fit) + Complexity penalty

۲- معیار BIC :

$$BIC = -2 \log g(y, \hat{\theta}(y)) + (\log n) \cdot p$$

تعداد متغیرها

تعداد پارامترها

اطلاعات درستی + حجم جریمه

۴- عامل بیز : که در تست خاتم احمدیان گفته شد

5 - DIC :

Deviance : $D(\theta) = -2 \log L(\theta | y)$

فرض لندہ با ~~...~~ تابع زبان مربع خطا (مثال) براؤن کرینر ہستم، در این صورت

$\theta_1, \dots, \theta_m \sim \pi(\theta | y)$ سب سے باری می ہستم

$\theta_1, \dots, \theta_m \sim \pi(\theta | y) \Rightarrow \hat{\theta}_B = E(\theta | y) = \frac{1}{m} \sum_{i=1}^m \theta_i = \bar{\theta}$

حال مقدار زیر را ہم میں لندہ

1) $D_i = -2 \log L(\theta_i | y) = D(\theta_i)$

$\hookrightarrow \overline{D(\theta)} = \frac{1}{m} \sum_{i=1}^m D_i$

2) $D(\bar{\theta}) = -2 \log L(\bar{\theta} | y)$

3) $P_D = \overline{D(\theta)} - D(\bar{\theta})$

$\Rightarrow DIC = D(\bar{\theta}) + 2P_D$

حفظ (مصرح) \rightarrow هر چه مدل سخته تر باشه این مقدار بزرگتر می شه و ما می خواهیم مدل ساده تر باشه
مثلاً $-2 \log L(\theta)$ این مقدار θ جایی $\hat{\theta} = \hat{\theta}_B$ که مقدار ستم

این جمله می گن مقدار خوب میسین
انجام می دهیم

نقطه: این BIC, AIC و DIC در fit کردن توزیع هم کاربرد دارد.

چون در واقع ما در مدل سازی می گیم مثلاً $Y_i \sim N(\mu_i, \sigma^2)$ و می گیم $\mu_i = X_i \beta$ و با حال

مثلاً $Y_i \sim t(\mu_i, \sigma^2)$ و $\mu_i = X_i \beta$ و بعد AIC, BIC, DIC این مدل ها را مقایسه می کنیم

خب توی fit کردن توزیع به معنی داره هم، همین کارایی کنیم مثلاً $\mu_i = X_i \beta$

و می گیم مثلاً $\mu_i = 4$!

10

15

20

25