Beginning of Jan.19, 2021

Theorem: (2.13) Central Limit Theorem, CLT

Let $X_1, ..., X_n$ be i.i.d. with $\mathbb{E}|X_1| < \infty$ and $0 < \mathrm{var}(X_1) < \infty$. Then for any $t \in \overline{\mathbb{R}}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}} \leqslant t\right) = \mathbb{P}(Z \leqslant t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-s^2/2}\, \mathrm{d}s,$$

where $\mu := \mathbb{E}X_1$ and $\sigma^2 := \mathrm{var}(x_1)$. In particular, each quotient $(X_1 + ... + X_n - n\mu)/(\sigma\sqrt{n})$ does have mean 0 and variance 1.

Theorem: (2.30) Berry-Esseén Theorem for CLT

Assume in addition that $\mathbb{E}|X_1|^3 < \infty$. Then

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(\frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}} \leqslant t\right) - \mathbb{P}(Z \leqslant t)\right| \leqslant \frac{\mathbb{E}|X_1|^3}{\sigma^3\sqrt{n}},$$

so in particular if $\mathbb{E}X_1 = 0$ and $\mathrm{var}(X_1) = 1$, we have

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}\left(\frac{X_1 + ... + X_n}{\sqrt{n}} \leqslant t\right) - \mathbb{P}(Z \leqslant t)\right| \leqslant \frac{\mathbb{E}|X_1|^3}{\sqrt{n}}.$$

Chapter 1

Exponential Families

A general question in statistics is to *fit a parameter to some given data*, for example, to find the unknown mean of a Gaussian sample.

An exponential family is some family of PDF or PMFs that depends on a parameter $w \in \mathbb{R}^k$ for some $k \geqslant 1$. More formally,

Definition: (3.1) Exponential Families

Let n, k be positive integers and let μ be a measure on $\mathbb{R}^n$. Let $t_1, \ldots, t_k : \mathbb{R}^n \to \mathbb{R}$, and let $h : \mathbb{R}^n \to [0, \infty]$ not identically zero. For any $w = (w_1, \ldots, w_k) \in \mathbb{R}^k$, define

$$a(w) := \log \int_{\mathbb{R}^n} h(x) \exp\big(\sum_{i=1}^{k} w_i t_i(x) \big) \, \mathrm{d}\mu(x).$$

The set $\{w \in \mathbb{R}^k : a(w) < \infty\}$ is called the **natural parameter space**. On this set, the functions

$$f_w(x) := h(x) \exp\left(\sum_{i=1}^{k} w_i t_i(x) - a(w) \right) \qquad \text{for all } x \in \mathbb{R}^n$$

satisfy

$$\begin{aligned} \int_{\mathbb{R}^n} f_w(x) \, \mathrm{d}x &= \int_{\mathbb{R}^n} h(x) \frac{\exp\big(\sum_{i=1}^{k} w_i t_i(x) \big)}{\int_{\mathbb{R}^n} h(x) \exp(\sum_{i=1}^{k} w_i t_i(x)) \, \mathrm{d}\mu(x)} \, \mathrm{d}\mu(x) \\ &= \frac{\int_{\mathbb{R}^n} h(x) \exp() \, \mathrm{d}\mu(x)}{\int_{\mathbb{R}^n} h(x) \exp() \, \mathrm{d}\mu(x)} = 1. \end{aligned}$$

Informally, the f_w's can be interpreted as probability density functions with respect to the measure μ. Then, the set of functions $\{f_w : a(w) < \infty\}$ is called a ***k*-parameter exponential family in canonical form**. (*We interpret f_w as a PDF or PMF according to μ the measure.*)

More generally, let $\Theta \subset \mathbb{R}^k$ and let $w : \Theta \to \mathbb{R}^k$. We define a ***k*-parameter exponential family** to be the set of functions $\{f_\theta : \theta \in \Theta, a(w(\theta)) < \infty\}$ where

$$f_\theta(x) := h(x) \exp\left(\sum_{i=1}^{k} w_i(\theta) t_i(x) - a(w(\theta)) \right) \qquad \text{for all } x \in \mathbb{R}^n.$$

Example: (3.3) Writing Gaussians as an Exponential Family. Consider Gaussians with mean $\mu < \infty$ and standard deviation $\sigma > 0$. Then the PDF is given by

$$\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{\mu x}{\sigma^2} - \frac{x^2}{2\sigma^2} - \left(\frac{\mu^2}{2\sigma^2} + \log\sigma\right)\right). \tag{1}$$

If we write $\theta = (\theta_1, \theta_2) := (\mu, \sigma^2) \in \mathbb{R}^2$ and define

$$t_1(x) := x, \qquad t_2(x) := x^2,$$
$$w_1(\theta) := \frac{\theta_1}{\theta_2} = \frac{\mu}{\sigma^2}, \qquad w_2(\theta) := -\frac{1}{2\theta_2} = -\frac{1}{2\sigma^2},$$
$$a(w(\theta)) := \frac{\theta_1^2}{2\theta_2} + \frac{1}{2}\log\theta_2 = \frac{\mu^2}{2\sigma^2} + \log\sigma,$$

and $h(x) := 1/\sqrt{2\pi}$, then (1) becomes

$$h(x)\exp\left(w_1(\theta)t_1(x) + w_2(\theta)t_2(x) - a(w(\theta))\right) \qquad \text{for all } x \in \mathbb{R}.$$

Let $\Theta := \mathbb{R} \times (0, \infty)$, and for $\theta \in \Theta$ we define

$$f_\theta(x) := h(x)\exp\left(\sum_{i=1}^{2} w_i(\theta)t_i(x) - a(w(\theta))\right) \qquad \text{for all } x \in \mathbb{R}.$$

From this we see that $\{f_\theta : \theta \in \Theta\}$ is a two parameter exponential family and that the Gaussians can be expressed by an exponential family.