- **定义5.1** （【0】定义2.6.1） 设$\mathcal{F}=\{f(\vec{x}|\vec{\theta}):\vec{\theta}\in\Theta\}$ 是一个参数统计模型，$\Theta$是参数空间，若其概率密度/质量函数可表达成如下形式：

$$f(\vec{x}|\vec{\theta})=C(\vec{\theta})\exp\left\{\sum_{i=1}^{k}Q_i(\vec{\theta})T_i(\vec{x})\right\}h(\vec{x}), \tag{1}$$

则称此分布族为**指数型分布族**（简称**指数族**，exponential family）. 其中$k$为正整数，$C(\vec{\theta})>0$和$Q_i(\vec{\theta})(i=1,\cdots,k)$ 都是定义在参数空间$\Theta$上的函数，$h(\vec{x})>0$和$T_i(\vec{x})(i=1,\cdots,k)$都是定义在样本空间$\mathcal{X}$ 上的函数.

注1 $f$的支撑集，也即集合$\{\vec{x}\in\mathbb{R}^d:f(\vec{x}|\vec{\theta})>0\}$不依赖于参数$\vec{\theta}$.

注2 表达式(1)也可记为

$$f(\vec{x}|\vec{\theta}) = \exp\left\{\sum_{i=1}^{k} Q_i(\vec{\theta}) T_i(\vec{x}) + A(\vec{\theta}) + B(\vec{x})\right\}.$$

注3 通常一维情形下的基本形式有$f(x)$, $g(\theta)$, $[f(x)]^{g(\theta)}$, $[g(\theta)]^{f(x)}$等.

注4 常见指数分布族：$Poisson(\lambda)$, $Binomial(n,p)$, $Gamma(\alpha,\beta)$和$N(\mu,\sigma^2)$等.

注5 验证一个参数分布族是否是指数族，等价于证明该分布族的概率密度/质量函数可表达成(1)，同时明确指出函数$C(\vec{\theta}) > 0$, $h(\vec{x})$, $Q_i(\vec{\theta})$以及$T_i(\vec{x}), i = 1, \cdots, k$。

# 举例说明

## Example (5.1)

[【0】例2.6.2] 设$X_1, \ldots, X_n \quad i.i.d. \sim Gamma(\alpha, \beta)$，p.d.f.

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\} \mathbf{1}_{(0,\infty)}(x)，\alpha, \beta > 0均未知，$$

证明其样本分布族为指数族。

## Example (5.2)

[【0】例2.6.3] 证明$Binomial(n, p), 0 < p < 1$未知，是指数族。

注 自行验证常见四类分布族以及各自的样本分布族是指数族。

- **反例** 证明一个参数分布族不是指数族，通常通过证明其概率密度函数的支撑集依赖于参数$\vec{\theta}$得以证明。

## Example (5.3)

[【0】例2.6.5] 均匀分布族$\{U(0, \theta), \theta > 0\}$不是指数族。

- **定义5.2** 如果指数族具有如下形式：

$$f(\vec{x}|\vec{\eta}) = C(\vec{\eta}) \exp\left\{\sum_{i=1}^{k} \eta_i T_i(\vec{x})\right\} h(\vec{x}),$$

则称它为**指数族的自然形式**（natural form）．此时集合

$$\left\{\vec{\eta} = (\eta_1, \ldots, \eta_k)^T : 0 < \int_{\mathcal{X}} \exp\{\textstyle\sum_{i=1}^{k} \eta_i T_i(\vec{x})\} h(\vec{x}) dx < \infty\right\}$$

称为**自然参数空间**（natural parametric space）．

## Example (5.4)

四类常见指数族的自然参数及其自然参数空间

| 分布 | 参数 | 自然参数$\vec{\eta}$ | 参数空间 |
|---|---|---|---|
| $N(\mu, \sigma^2)$ | $\mu, \sigma^2$ | $(\frac{\mu}{\sigma^2}, \frac{1}{2\sigma^2})$ | $\{(\eta_1, \eta_2) : -\infty < \eta_1 < \infty, \eta_2 > 0\}$ |
| $Gamma(\alpha, \beta)$ | $\alpha, \beta$ | $(\alpha - 1, -\beta)$ | $\{(\eta_1, \eta_2) : \eta_1 > -1, \eta_2 < 0\}$ |
| $Binomial(n, p)$ | $p$ | $\ln \frac{p}{1-p}$ | $\{\eta : -\infty < \eta < \infty\}$ |
| $Poisson(\lambda)$ | $\lambda$ | $\ln \lambda$ | $\{\eta : -\infty < \eta < \infty\}$ |

**Theorem (5.1)**

*[【0】定理2.6.2]* 设指数族的自然形式中，自然参数空间有内点，其内点集为$\Theta_0$，设$g(x)$为任一实函数，使得积分

$$G(\overrightarrow{\eta}) = \int g(x) \exp\{\sum_{i=1}^{k} \eta_i T_i(\overrightarrow{x})\} h(\overrightarrow{x}) d\overrightarrow{x}$$

在$\Theta_0$内存在有限，则$G(\overrightarrow{\eta})$的任意阶偏导数在$\Theta_0$内存在且可在积分号下求得，即

$$\frac{\partial^m}{\partial \eta_1^{m_1} \cdots \partial \eta_k^{m_k}} G(\overrightarrow{\eta}) = \int \frac{\partial^m}{\partial \eta_1^{m_1} \cdots \partial \eta_k^{m_k}} g(x) \exp\{\sum_{i=1}^{k} \eta_i T_i(\overrightarrow{x})\} h(\overrightarrow{x}) d\overrightarrow{x}$$

其中$\sum_{i=1}^{k} m_i = m$，即积分和求偏导顺序可交换。

# 作业汇总

1. 证明具有如下概率密度函数的分布族不是指数族：

$$f(x|\mu) = \frac{1}{4}\exp\left\{-\frac{1}{2}|x-\mu|\right\}, -\infty < x < \infty,$$

这里$-\infty < \mu < \infty$.

2. 习题2：Ex. 39, 40.