STAT3602 Statistical Inference

Example Class 6

LIU Chen

Department of Statistics and Actuarial Science, HKU

07/10/2020

Outline

Definition: Exponential Families

A family of distributions $\{\mathbb{P}_\theta : \theta \in \Theta\}$ on a sample space $\mathcal{S}$ (free of θ) is an exponential family if its probability functions are of the form

$$f(x \mid \theta) \propto \begin{cases} h(x) \exp\left\{\sum_{j=1}^{k} \pi_j(\theta) t_j(x)\right\}, & x \in \mathcal{S} \\ 0, & \text{otherwise} \end{cases}$$

Without loss of generality, assume $h(x) > 0$ for $x \in \mathcal{S}$.

Natural parameters and natural statistics I

Suppose $\boldsymbol{X}$ has probability function of exponential family form with natural statistic $\boldsymbol{T} = (t_1(\boldsymbol{X}), \ldots, t_k(\boldsymbol{X}))$ and natural parameter $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k)$.

- **natural parameter space.**
  The set $\Pi = \{\boldsymbol{\pi} = (\pi_1(\theta), \pi_2(\theta), \ldots, \pi_k(\theta)) : \theta \in \Theta\} \subset \mathbb{R}^k$.
- **natural parameter.** The vector $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k) \in \Pi$.
- **natural statistic.** for a random variable $\boldsymbol{X}$ drawn from $f(\boldsymbol{x} \mid \theta)$, the natural statistic is $(t_1(\boldsymbol{X})), \ldots, t_k(\boldsymbol{X}))$.

The distributional Properties of natural statistic

Theorem 3.3.1. Then $\boldsymbol{T}$ has probability function also of exponential family form with the same natural parameter.
Theorem 3.3.2. The joint distribution of a partial list of the natural statistics conditioned on the rest is in exponential family form, with the corresponding natural parameters.

Outline

(d) Beta Distribution:

$$f(x \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in (0,1), \alpha, \beta > 0.$$

(g) Negative binomial (Pascal distribution):

$$p(x; r, p) = \binom{r+x-1}{r-1} p^r (1-p)^x, x = 1, 2, \ldots;$$

with $p \in (0,1), r = \{1, 2, \ldots\}$.

Exercise 1: Solution I

(d) Beta Distribution:
$\theta = (\alpha, \beta)$ and the density is

$$f(x \mid \theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} e^{(\alpha-1)\log x + (\beta-1)\log(1-x)}$$

where $x \in (0,1), \alpha, \beta > 0$.
So, Beta(α, β) belongs to exponential family with

$$\begin{aligned}
c(\theta) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}, \\
h(x) &= 1, \\
\pi_1(\theta) &= \beta - 1, \\
\pi_2(\theta) &= \alpha - 1.
\end{aligned}$$

Exercise 1: Solution II

Then, the pdf could be written as

$$f(x \mid \theta) = \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)} e^{\pi_2 \log x + \pi_1 \log(1-x)},$$

where $\beta = \pi_1 + 1, \alpha = \pi_2 + 1$.

- The natural parameter is $\pi = (\pi_1, \pi_2) = (\beta - 1, \alpha - 1)$.
- The natural parameter space is $\Pi = \{\pi : c(\pi) > 0\} = \left\{\pi : \frac{\Gamma(\pi_1+\pi_2+2)}{\Gamma(\pi_1+1)\Gamma(\pi_2+1)} > 0\right\} = (-1, \infty) \times (-1, \infty)$.

The joint density of i.i.d random variables, $x_1, \ldots, x_n$ is

$$f(x_1, \ldots, x_n \mid \alpha, \beta) = \prod_{i=1}^{n} \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)} e^{\pi_2 \log x_i + \pi_1 \log(1-x_i)}$$

$$= \left[\prod_{i=1}^{n} \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)}\right] e^{\pi_2 \sum_{i=1}^{n} \log x_i + \pi_1 \sum_{i=1}^{n} \log(1-X_i)}.$$

The natural statistic is $\sum_{i=1}^{n} \log X_i$ and $\sum_{i=1}^{n} \log(1 - X_i)$.

Exercise 1: Solution III

(g) Negative binomial (Pascal distribution):
The probability mass function of a negative binomial distribution is

$$p(x; r, p) = \binom{r+x-1}{r-1} p^r (1-p)^x, x = 1, 2, \ldots; p \in (0, 1), r = \{1, 2, \ldots$$

which can be expressed as

$$\begin{aligned} p(x; r, p) &= \frac{(x+r-1)!}{(r-1)!x!} p^r (1-p)^x \\ &= \frac{(x+r-1)(x+r-2)\cdots(x+1)}{(r-1)!} p^r (1-p)^x \\ &= p^r \left[\gamma_{r-1} x^{r-1} + \gamma_{r-2} x^{r-2} + \cdots + \gamma_0\right] \exp\{\log(1-p)x\} \end{aligned}$$

where $\gamma_0, \ldots, \gamma_{r-1}$ are positive terms involving r only.
It is obvious that the polynomial in x cannot be arranged into the desired form of $\exp\left\{\sum \pi_i(r)\tau_i(x)\right\}$. Therefore, a negative binomial

Exercise 1: Solution IV

distribution with both parameters unknown does not belong to the exponential family. (N.B. The negative binomial distribution with only one unknown parameter p belongs to the exponential family.)

Outline

Exercise 2

Find the general form of a conjugate prior density for θ in a Bayesian analysis of the one-parameter exponential family density

$$f(x;\theta) = c(\theta)h(x)\exp\{\theta t(x)\}, \quad x \in \mathbb{R}.$$

[If the posterior distribution $p(\theta \mid x)$ and the prior $p(\theta)$ have the same family of distribution, the prior is called a conjugate prior for the likelihood function.]

Exercise 2: Solution I

Rewrite the density of X as

$$f(x;\theta)=h(x)\exp\{\theta t(x)-A(\theta)\},$$

with $A(\theta)=-\log c(\theta)$.
The posterior density of Θ is

$$\begin{aligned}\pi(\theta\mid x)&\propto\pi(\theta)\prod_{i=1}^{n}f\left(x_i\mid\theta\right)\\&=\pi(\theta)\prod_{i=1}^{n}\left\{h\left(x_i\right)\right\}\exp\left\{\theta\sum_{i=1}^{n}t\left(x_i\right)-nA(\theta)\right\}\\&\propto\pi(\theta)\exp\left\{\theta\sum_{i=1}^{n}t\left(x_i\right)-nA(\theta)\right\}\end{aligned}$$

Exercise 2: Solution II

For π to be a conjugate prior, we must have

$$\pi(\theta) \propto \exp\{\eta\theta - vA(\theta)\},$$

i.e. the distribution of Θ belongs to the exponential family with natural statistic $(\theta, A(\theta))$. If we denote the prior distribution as $\Pi(\eta, v)$, the posterior of Θ given $\boldsymbol{X} = (X_1, \ldots, X_n)$ is $\Pi\left(\eta + \sum t\left(x_i\right), v + n\right)$.

Outline

Exercise 3 I

A curved exponential family is a k-parameter exponential family (the dimension of the natural parameter is k) where the dimension of the vector θ is $d < k$. Let $Y_1, \ldots, Y_n$ be independent, identically distributed $N\left(\mu, \mu^2\right)$ variables. Show that this model is an example of a curved exponential family.

Exercise 3: Solution I

The dimension of the parameter vector μ is 1, i.e. $d = 1$.
The probability function of $(Y_1, \ldots, Y_n)$ is

$$
\begin{aligned}
f(y_1, \ldots, y_n \mid \mu) &= \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\mu} \exp\left\{ -\frac{(y_i - \mu)^2}{2\mu^2} \right\} \right) \\
&\propto \exp\left\{ -\frac{1}{2\mu^2} \sum_{i=1}^{n} y_i^2 + \frac{1}{\mu} \sum_{i=1}^{n} y_i \right\}.
\end{aligned}
$$

The joint pdf is of the form of the exponential family with natural parameter $\left(-\frac{1}{2\mu^2}, \frac{1}{\mu}\right)$ and natural statistic $\left(\sum_{i=1}^{n} y_i^2, \sum_{i=1}^{n} y_i\right)$. The natural parameter are two dimensional vectors. Therefore, such model belongs to the curved exponential family.

Outline

Exercise 4 I

Let X_1 and X_2 be independent normal random variables with means 0, and variances σ_1^2 and σ_2^2 respectively, so that the density function of X_i is

$$f\left(x \mid \sigma_i^2\right)=\frac{1}{\sqrt{2 \pi \sigma_i^2}} \exp \left\{-\frac{x^2}{2 \sigma_i^2}\right\}, \quad x \in R, i=1,2.$$

1. Define $\xi=\frac{1}{2}\left(\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\right)$.
   Show that the joint distribution of X_1, X_2 has the form of an exponential family with natural parameter $\left(\xi,-1 /\left(2 \sigma_2^2\right)\right)$ and natural statistic $\left(X_1^2, X_1^2+X_2^2\right)$
2. Show that the conditional distribution of X_1^2 given $X_1^2+X_2^2$ depends only on the parameter ξ.

Exercise 4: Solution I

1. The joint distribution of X_1 and X_2 is:

$$
\begin{aligned}
f\left(x_1, x_2 \mid \sigma_1, \sigma_2\right) &\propto \exp\left\{-\frac{x_1^2}{2\sigma_1^2} - \frac{x_2^2}{2\sigma_2^2}\right\} \\
&= \exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)x_1^2 + \left(-\frac{1}{2\sigma_2^2}\right)\left(x_1^2 + x_2^2\right)\right\},
\end{aligned}
$$

   which has the form of an exponential family with natural parameter $\left(\xi, -1/\left(2\sigma_2^2\right)\right)$ and natural statistic $\left(X_1^2, X_1^2 + X_2^2\right)$.

2. We use the property of an exponential family distribution, that the joint distribution of a partial list of the natural statistics conditioned on the rest is in exponential family form, with the corresponding natural parameters. Together with the results from (a), we are ready to conclude that $X_1^2 \mid X_1^2 + X_2^2$ depends only on the parameter ξ.