THE UNIVERSITY OF HONG KONG
DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

STAT3602 Statistical Inference

(2020-2021 First Semester)

Example Class 6 Solution[Exponential Families]

Date: 28/10/2020

1. Show whether the following distributions belong to the exponential family. If so, give its natural parameters and natural parameter space, and describe the natural statistic based on a random sample $x_1, \ldots, x_n$

   (a) Binomial Distribution:

   $$p(x \mid \theta) = \binom{k}{x} \theta^x (1-\theta)^{k-x}, \quad x = 0, \ldots, k; \theta \in (0,1).$$

   (b) Poisson Distribution:

   $$p(x \mid \theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \ldots, \theta > 0.$$

   (c) Gamma Distribution:

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

   (d) Beta Distribution:

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

   (e) Multinomial Distribution

   $$p(\boldsymbol{x}; \boldsymbol{\theta}) = \binom{N}{x_1, \ldots, x_{k-1}, \left(N - \sum_{i=1}^{k-1} x_i\right)} \theta_1^{x_1} \cdots \theta_{k-1}^{x_{k-1}} \left(1 - \sum_{i=1}^{k-1} \theta_i\right)^{N - \sum_{i=1}^{k-1} x_i},$$

   with $x_1, \ldots, x_{k-1}, \sum_{i=1}^{k-1} x_i \in \{0, 1, \ldots, N\}; \theta_1, \ldots, \theta_{k-1}, \sum_{i=1}^{k-1} \theta_i \in (0,1)$.

   (f) Uniform Distribution:

   $$f(x; \theta) = \theta^{-1}, \quad 0 < x < \theta; \theta \in (0, \infty).$$

   (g) Negative binomial (Pascal distribution):

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

(a) Binomial Distribution:

$$p(x \mid \theta)=\binom{k}{x}\theta^{x}(1-\theta)^{k-x}=\binom{k}{x}\left(\frac{\theta}{1-\theta}\right)^{x}(1-\theta)^{k}=\binom{k}{x}e^{x\log\frac{\theta}{1-\theta}}(1-\theta)^{k}$$

Hence, $h(x)=\binom{k}{x}, c(\theta)=(1-\theta)^{k}$ and $\pi(\theta)=\log\frac{\theta}{1-\theta}$. Thus, the pmf of x given θ is

$$p(x \mid \theta)=\binom{k}{x}\frac{1}{(e^{\pi}+1)^{k}}e^{x\pi},\ \text{where } \theta=\frac{1}{e^{-\pi}+1}.$$

Natural parameter is $\pi=\log\left(\frac{\theta}{1-\theta}\right)$, and the natural parameter space is $\Pi=\{\pi : c(\pi)>0\}=\left\{\pi : \frac{1}{(e^{\pi}+1)^{k}}>0\right\}=(-\infty,\infty)$.

The joint p.m.f of i.i.d binomial distributed random variables $x_1,\ldots,x_n$ is

$$p\left(x_{1},\ldots,x_{n} \mid \theta\right)=\prod_{i=1}^{n}\binom{k}{x_{i}}e^{x_{i}\pi}\frac{1}{(e^{\pi}+1)^{k}}=\left[\prod_{i=1}^{n}\binom{k}{x_{i}}\right]e^{\pi\sum_{i=1}^{n}x_{i}}\frac{1}{(e^{\pi}+1)^{nk}}$$

The natural statistic is $\sum_{i=1}^{n}X_i$.

(b) Poisson Distribution:

$$p(x \mid \theta)=\frac{\theta^{x}e^{-\theta}}{x!}=\frac{1}{x!}e^{-\theta}e^{-x\log\theta}\quad x=0,1,\ldots,\theta>0.$$

Thus, $h(x)=\frac{1}{x!}, c(\theta)=e^{-\theta}$, and $\pi(\theta)=\log\theta$. Thus, the pmf of x given θ is

$$p(x \mid \theta)=\frac{1}{x!}e^{-\theta}e^{-x\log\theta}=\frac{1}{x!}e^{-e^{\pi}}e^{-x\pi},\quad \text{with } \theta=e^{\pi}.$$

Thus, the natural parameter is $\pi(\theta)=\log\theta$. The natural parameter space is $\Pi=\{\pi : c(\pi)>0\}=\{\pi:\exp\{-ne^{\pi}\}>0\}=(-\infty,\infty)$

The joint p.m.f of $x_1,\ldots,x_n$ is

$$p\left(x_{1},\ldots,x_{n} \mid \theta\right)=\prod_{i=1}^{n}\frac{1}{x_{i}!}e^{-e^{\pi}}e^{-x_{i}\pi}=\left[\prod_{i=1}^{n}\frac{1}{x_{i}!}\right]e^{-ne^{\pi}}e^{-\pi\sum_{i=1}^{n}x_{i}}.$$

Thus, the natural statistic is $\sum_{i=1}^{n}X_i$.

(c) Gamma Distribution:

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

$\theta=(\alpha,\beta)$, and the density is

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

. So, Gamma (α,β) belongs to exponential families with $c(\theta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}, h(x)=1$ and $\pi_1(\theta)=-\beta, \pi_2(\theta)=\alpha-1$. Thus, the pdf of x given θ could be written as

$$f(x \mid \alpha,\beta)=\frac{\pi_{1}^{\pi_{2}+1}}{\Gamma\left(\pi_{2}+1\right)}e^{\pi_{1}x+\pi_{2}\log x},\ \text{with } \beta=-\pi_{1},\alpha=\pi_{2}+1$$

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

The joint density of $x_1,\ldots,x_n$ is

$$f\left(x_{1},\ldots,x_{n} \mid \alpha,\beta\right)=\prod_{i=1}^{n}\frac{\pi_{1}^{\pi_{2}+1}}{\Gamma\left(\pi_{2}+1\right)}e^{\pi_{1}x_{i}+\pi_{2}\log x_{i}}=\left[\prod_{i=1}^{n}\frac{\pi_{1}^{\pi_{2}+1}}{\Gamma\left(\pi_{2}+1\right)}\right]e^{\pi_{1}\sum_{i=1}^{n}x_{i}+\pi_{2}\sum_{i=1}^{n}\log x_{i}}.$$

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

(d) Beta Distribution:

$\theta = (\alpha, \beta)$ and the density is

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

So, Beta(α, β) belongs to exponential family with $c(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}, h(x) = 1$ and $\pi_1(\theta) = \beta - 1, \pi_2(\theta) = \alpha - 1$. Then, the pdf could be written as

$$f(x \mid \theta) = \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)} e^{\pi_2 \log x + \pi_1 \log(1-x)} \text{ with } \quad \beta = \pi_1 + 1, \alpha = \pi_2 + 1.$$

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

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

$$\begin{aligned} f(x_1, \ldots, x_n \mid \alpha, \beta) &= \prod_{i=1}^{n} \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)} e^{\pi_2 \log x_i + \pi_1 \log(1-x_i)} \\ &= \left[\prod_{i=1}^{n} \frac{\Gamma(\pi_1 + \pi_2 + 2)}{\Gamma(\pi_1 + 1)\Gamma(\pi_2 + 1)}\right] e^{\pi_2 \sum_{i=1}^{n} \log x_i + \pi_1 \sum_{i=1}^{n} \log(1-X_i)}. \end{aligned}$$

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

(e) Multinomial Distribution

$$p(\boldsymbol{x}; \boldsymbol{\theta}) = \binom{N}{x_1, \ldots, x_{k-1}, \left(N - \sum_{i=1}^{k-1} x_i\right)} \theta_1^{x_1} \cdots \theta_{k-1}^{x_{k-1}} \left(1 - \sum_{i=1}^{k-1} \theta_i\right)^{N - \sum_{i=1}^{k-1} x_i},$$

with $x_1, \ldots, x_{k-1}, \sum_{i=1}^{k-1} x_i \in \{0, 1, \ldots, N\}; \theta_1, \ldots, \theta_{k-1}, \sum_{i=1}^{k-1} \theta_i \in (0, 1)$.

For any allowed combination of $\theta_i, i = 1, \ldots, k-1$, the joint pmf of $x_1, \ldots, x_{k-1}$ can be written as

$$\begin{aligned} p(\boldsymbol{x}; \boldsymbol{\theta}) &= \binom{N}{x_1, \ldots, x_{k-1}, \left(N - \sum_{i=1}^{k-1} x_i\right)} \theta_1^{x_1} \cdots \theta_{k-1}^{x_{k-1}} \left(1 - \sum_{i=1}^{k-1} \theta_i\right)^{N - \sum_{i=1}^{k-1} x_i} \\ &= h(\boldsymbol{x}) \exp\left\{\sum_{i=1}^{k-1} x_i \log\left(\frac{\theta_i}{1 - \sum_{j=1}^{k-1} \theta_j}\right)\right\} \left(1 - \sum_{j=1}^{k-1} \theta_j\right)^N \\ &= h(\boldsymbol{x}) \exp\left\{\sum_{i=1}^{k-1} \pi_i(\boldsymbol{\theta}) \tau_i(\boldsymbol{x})\right\} c(\boldsymbol{\theta}), \end{aligned}$$

where $h(\boldsymbol{x}) = \left(x_1, \ldots, x_{k-1}, \left(N - \sum_{i=1}^{k-1} x_i\right)\right), \pi_i = \log\left(\frac{\theta_i}{1 - \sum_{j=1}^{k-1} \theta_j}\right), \tau_i(\boldsymbol{x}) = x_i$ and $c(\boldsymbol{\theta}) = \left(1 - \sum_{j=1}^{k-1} \theta_j\right)^N$. So, the multinomial distribution also forms an exponential family, The natural parameter is $\pi(\theta) = \left(\log\left[\theta_1 / \left(1 - \sum_{i=1}^{k-1} \theta_i\right)\right], \cdots, \log\left[\theta_{k-1} / \left(1 - \sum_{i=1}^{k-1} \theta_i\right)\right]\right)'$. and the natural parameter space is

$$\begin{aligned} \Pi &= \left\{(\pi_1, \cdots, \pi_{k-1})' : 0 < \sum_x h(x) \exp\left\{\sum_{i=1}^{k-1} x_i \pi_i\right\} < \infty\right\} \\ &= (-\infty, \infty)^{k-1}. \end{aligned}$$

The joint pmf of $X_1, \ldots, X_n$ is

$$\prod_{j=1}^{n} h(\boldsymbol{x}_j) \times \exp\left\{\sum_{i=1}^{k-1} \pi_i \sum_{j=1}^{n} \tau_i(\boldsymbol{x}_j)\right\} \times J(\boldsymbol{\pi})^{-n}$$

Therefore, the natural statistic is $t = \left(\sum_{j=1}^{n} x_{j1}, \ldots, \sum_{j=1}^{n} x_{j,k-1}\right)'$

(f) Uniform Distribution:

$$f(x;\theta)=\theta^{-1}, \quad 0<x<\theta;\theta\in(0,\infty).$$

It does not belong to the exponential family as the sample space $(0,\theta)$ depends on the unknown parameter θ.

(g) Negative binomial (Pascal distribution):

The probability mass function of a negative binomial distribution is

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

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 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.)

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.]

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

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

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 $\Pi\left(\eta+\sum t\left(x_i\right), v+n\right)$.

3. 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.

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.

4. 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

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

(a) 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 $\left(X_1^2, X_1^2 + X_2^2\right)$

(b) Show that the conditional distribution of X_1^2 given $X_1^2 + X_2^2$ depends only on the parameter ξ.

(a) The joint distribution of X_1 and X_2 is:

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

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

(b) 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.