Stat 512
Homework key 3

October 14, 2015

REGULAR PROBLEMS

1. Suppose random variable $X \sim \mathcal{N}(\mu, \sigma^2)$. Show that $Y = e^X$ has a log normal distribution.

   Solution. Since $h(x) = \exp(x)$ is monotonic increasing ($h'(x) = h(x) = \exp(x) > 0$), the transformation follows from results given in class. In particular, since $X = \log Y \equiv h^{-1}(Y)$, we have that the pdf of Y is

$$
\begin{aligned}
g(y) &= \phi(h^{-1}(y))\frac{d}{dy}h^{-1}(y) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{1}{2\sigma^2}(h^{-1}(y)-\mu)^2\right)\mathbf{1}[h^{-1}(y)\int_{-\infty}^{\infty}\mathbb{R}]\frac{d}{dy}h^{-1}(y) \\
&= \frac{1}{y\sqrt{2\pi\sigma^2}}\exp\left(\frac{1}{2\sigma^2}(\log y-\mu)^2\right)\mathbf{1}[y>0].
\end{aligned}
$$

   We see that Y has the pdf of the log normal family and hence is distributed log normal $\mathcal{LN}(\mu, \sigma^2)$.

2. Suppose we transform a random variable X according to $Y = aX + b$. Show that the following parametric distribution families are closed under such a transformation (i.e., that they are special cases of "location-scale families of distributions").

   (a) $X \sim \mathcal{U}(\alpha, \beta)$ has $Y \sim \mathcal{U}(a\alpha + b, a\beta + b)$.

   (b) $X \sim \Gamma(\alpha, \beta, A)$ has $Y \sim \Gamma(\alpha, a\beta, aA + b)$.

   (c) $X \sim \mathcal{N}(\mu, \sigma^2)$ has $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

   Solution. We first note that there are no explicit restrictions placed on a in this problem. However, in order for the families to be closed under the transformation $Y = aX + b$, we will sometimes impose additional restrictions on the scale parameter. Most generally, we require that $a \neq 0$ in each of (a)-(c). If not, $Y \equiv b$, so $P(Y = y) = \mathbf{1}_{\{b\}}(y)$. This is a degenerate probability distribution at b, and clearly not a member of the given parametric families. Having now handled the trivial case, we provide a useful result for location-scale families.

   Lemma 1. If X has cdf F, pdf f and support $\mathcal{S}$, then (for $a \neq 0$, $b \int_{-\infty}^{\infty} \mathbb{R}$) $Y \equiv h(X) = aX + b$ has pdf given by

$$
g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right)\mathbf{1}_{\mathcal{S}}\left(\frac{y-b}{a}\right).
$$

Proof. This result follows as a corollary of the monotone transformation formulas derived in class. Alternatively, note that

$$\begin{aligned}
Pr(Y \leq y) &= Pr(h(X) \leq y) \\
&= Pr(aX + b \leq y) \\
&= Pr\left(X \leq \frac{y-b}{a}\right)\mathbf{1}[a > 0] + Pr\left(X \geq \frac{y-b}{a}\right)\mathbf{1}[a < 0] \\
&= F\left(\frac{y-b}{a}\right)\mathbf{1}[a > 0] + \left[1 - F\left(\frac{y-b}{a}\right)\right]\mathbf{1}[a < 0] \\
&\equiv G(y).
\end{aligned}$$

Differentiating, we find

$$g(y) = \frac{dG}{dy}(y) = \frac{1}{a}f\left(\frac{y-b}{a}\right)\mathbf{1}[a > 0] + \left[-\frac{1}{a}f\left(\frac{y-b}{a}\right)\right]\mathbf{1}[a < 0] = \frac{1}{|a|}f\left(\frac{y-b}{a}\right),$$

for all y such that $(y-b)/a \int_{-\infty}^{\infty} \mathcal{S}$, and zero elsewhere.

(a) We have that $X \sim \mathcal{U}(\alpha, \beta)$, so $\alpha < \beta$ and

$$f(x) = \frac{\mathbf{1}[\alpha < x < \beta]}{\beta - \alpha}.$$

By Lemma 1, for $a > 0$ and $Y = aX + b$, we have

$$\begin{aligned}
g(y) &= \frac{1}{|a|}f\left(\frac{y-b}{a}\right) \\
&= \frac{1}{a}\frac{\mathbf{1}[\alpha < (y-b)/a < \beta]}{\beta - \alpha} \\
&= \frac{1}{a}\frac{\mathbf{1}[\alpha < (y-b)/a < \beta]}{\beta - \alpha} \\
&= \frac{\mathbf{1}[a\alpha + b < y < a\beta + b]}{a\beta + b - (a\alpha + b)},
\end{aligned}$$

which is the pdf of the $\mathcal{U}(a\alpha + b, a\beta + b)$. Hence, the uniform distributions are closed under location-scale transformations in this case. We note, however, that for $a < 0$

$$\begin{aligned}
g(y) &= \frac{1}{|a|}f\left(\frac{y-b}{a}\right) \\
&= \frac{-1}{a}\frac{\mathbf{1}[\alpha < (y-b)/a < \beta]}{\beta - \alpha} \\
&= \frac{1}{a}\frac{\mathbf{1}[\alpha < (y-b)/a < \beta]}{\beta - \alpha} \\
&= \frac{\mathbf{1}[a\beta + b < y < a\alpha + b]}{a\alpha + b - (a\beta + b)}.
\end{aligned}$$

This is again the pdf of a uniform distribution, but this time with an additional reflection across the origin (i.e. $\mathcal{U}(a\beta + b, a\alpha + b)$).

(b) We have that $X \sim \Gamma(\alpha, \beta, A)$ for $\alpha, \beta > 0$. Hence the pdf of X is

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}(x - A)^{\alpha-1}e^{-(x-A)/\beta}\mathbf{1}[x > A].$$

By Lemma 1, for $a > 0$ and $Y = aX + b$, we have that the pdf of Y is

$$\begin{aligned} g(y) &= \frac{1}{|a|} f\left(\frac{y-b}{a}\right) \\ &= \frac{1}{a}\frac{1}{\Gamma(\alpha)\beta^\alpha}((y-b)/a - A)^{\alpha-1}e^{-((y-b)/a-A)/\beta}\mathbf{1}[(y-b)/a > A] \\ &= \frac{1}{\Gamma(\alpha)(a\beta)^\alpha}(y - (aA+b))^{\alpha-1}e^{-(y-(aA+b))/(a\beta)}\mathbf{1}[y > aA + b]. \end{aligned}$$

Hence Y has the pdf (and distribution) of $\Gamma(\alpha, a\beta, aA+b)$, demonstrating the appropriate closure of the family under the linear transformation. Again note that while we do attain a valid pdf in the case that $a < 0$, but the resultant scale parameter $a\beta$ will be negative, and the support will be similarly reflected. We have now seen how distributions such as the uniform and gamma families motivate the common restriction of location-scale families to positive scale parameters $a > 0$.

(c) We have that $X \sim \mathcal{N}(\mu, \sigma)$. Hence the pdf of X is

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{1}{2\sigma^2}(x-\mu)^2\right)\mathbf{1}[x \int_{-\infty}^{\infty} \mathbb{R}].$$

By Lemma 1, for $a \neq 0$ and $Y = aX + b$, we have that the pdf of Y is

$$\begin{aligned} g(y) &= \frac{1}{|a|}\phi\left(\frac{y-b}{a}\right) \\ &= \frac{1}{|a|}\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{1}{2\sigma^2}((y-b)/a - \mu)^2\right)\mathbf{1}[(y-b)/a \int_{-\infty}^{\infty} \mathbb{R}] \\ &= \frac{1}{\sqrt{2\pi|a|^2\sigma^2}} \exp\left(\frac{1}{2|a|^2\sigma^2}((y-(a\mu+b))^2\right)\mathbf{1}[y \int_{-\infty}^{\infty} \mathbb{R}]. \end{aligned}$$

Hence, Y has the pdf (and distribution) of a $\mathcal{N}(a\mu + b, a^2\sigma^2)$ distribution. The normal distributions are closed under the transformation $Y = aX + b$ for any $a \neq 0$. Note that in this case, $a < 0$ does not lead to the problem with the definition (notation) of the parametric family observed in (a) and (b).

(Note: Though seemingly mundane, a brief acknowledgment of the implicit restriction $a > 0$ is a key element of the solution for parts (a) and (b). Noting the degeneracy of the trivial case is of similar importance. Identifying and clearly communicating technical details and assumptions are important skills to develop for work in both applied and theoretical statistics.)

3. Let $\vec{X} = (X_1, \ldots, X_n)$ be a random vector in which the X_i are independently distributed with an exponential family distribution having density (probability mass function) of the form

$$f_{\vec{X}}(\vec{x}|\vec{\theta}) = h(\vec{x}) \exp\left[\sum_{i=1}^{p} T_i(\vec{x})\eta_i(\vec{\theta}) + A(\vec{\theta})\right].$$

For each of the following parametric distributions, show whether it does or does not belong to an exponential family by explicitly identifying $h(\vec{x})$, $\vec{\eta}(\vec{\theta})$, $\vec{T}(\vec{x})$, and $A(\vec{\theta})$ where possible. For every exponential family distribution, be sure to indicate the dimensionality of the statistic $\vec{T}(\vec{X})$ and whether the exponential family is curved. (Recall that the density (pmf) for a random vector of independent random variables is given by

$$f_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} f_{X_i}(x_i),$$

where f_{X_i} is the density (pmf) for the ith component of the random vector $\vec{X}$.

(a) Bernoulli distribution: $X_i \sim \mathcal{B}(1, p)$ with $\theta = p$.

Solution.

$$\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= \exp(\sum_{i=1}^{n} x_i \log \frac{p}{1-p} + n\log(1-p)) \\
&h(\overrightarrow{x}) = 1 \\
&\overrightarrow{T}(\overrightarrow{x}) = \sum_{i=1}^{n} x_i \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \log\left(\frac{p}{1-p}\right) \\
&p = 1 \\
&A(\overrightarrow{\theta}) = n\log(1-p) \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}$$

Since dimension of $\theta = p$, this family is not curved.

(b) Binomial distribution with varying replications: $X_i \sim \mathcal{B}(m_i, p)$ with $\theta = p$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \binom{m_i}{x_i} p^{x_i}(1-p)^{m_i - x_i} \\
&= \left[\prod_{i=1}^{n} \binom{m_i}{x_i}\right] \exp(\sum_{i=1}^{n} x_i \log \frac{p}{1-p} + \log(1-p) \sum_{i=1}^{n} m_i) \\
&h(\overrightarrow{x}) = \prod_{i=1}^{n} \binom{m_i}{x_i} \\
&\overrightarrow{T}(\overrightarrow{x}) = \sum_{i=1}^{n} x_i \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \log\left(\frac{p}{1-p}\right) \\
&p = 1 \\
&A(\overrightarrow{\theta}) = \log(1-p) \sum_{i=1}^{n} m_i \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(c) Geometric distribution: $X_i \sim Geom(p)$ with $\theta = p$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} p(1-p)^{x_i - 1} \\
&= \exp(\sum_{i=1}^{n} x_i \log(1-p) + n \log \frac{p}{1-p} \\
&h(\overrightarrow{x}) = 1 \\
&\overrightarrow{T}(\overrightarrow{x}) = \sum_{i=1}^{n} x_i \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \log(1-p) \\
&p = 1 \\
&A(\overrightarrow{\theta}) = n \log \frac{p}{1-p} \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(d) Negative binomial distribution with varying number of events: $X_i \sim NegB(r_i, p)$ with $\theta = p$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \binom{r_i + x_i - 1}{x_i} p^{r_i}(1-p)^{x_i} \\
&= \left[\prod_{i=1}^{n} \binom{r_i + x_i - 1}{x_i}\right] \exp(\sum_{i=1}^{n} x_i \log(1-p) + \sum_{i=1}^{n} r_i \log p) \quad h(\overrightarrow{x}) = \prod_{i=1}^{n} \binom{r_i + x_i - 1}{x_i} \\
&\overrightarrow{T}(\overrightarrow{x}) = \sum_{i=1}^{n} x_i \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \log(1-p) \\
&p = 1 \\
&A(\overrightarrow{\theta}) = \sum_{i=1}^{n} r_i \log p \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(e) Poisson distribution: $X_i \sim \mathcal{P}(\lambda)$ with $\theta = \lambda$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \\
&= \left[\prod_{i=1}^{n} \frac{1}{x_i!}\right] \exp(\sum_{i=1}^{n} x_i \log \lambda + -n\lambda) \\
&h(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{1}{x_i!} \\
&\overrightarrow{T}(\overrightarrow{x}) = \sum_{i=1}^{n} x_i \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \log \lambda \\
&p = 1 \\
&A(\overrightarrow{\theta}) = -n\lambda \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(f) Poisson distribution with varying time of follow-up: $X_i \sim \mathcal{P}(\lambda t_i)$ with $\theta = \lambda$ (t_i's are known constants).

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} e^{-\lambda t_i} \frac{(\lambda t_i)^{x_i}}{x_i!} \\
&= \left[\prod_{i=1}^{n} \frac{t_i^{x_i}}{x_i!}\right] \exp(\sum_{i=1}^{n} x_i \log \lambda + - \sum_{i=1}^{n} t_i \lambda) \\
h(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{t_i^{x_i}}{x_i!} \\
\overrightarrow{T}(\overrightarrow{x}) &= \sum_{i=1}^{n} x_i \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= \log \lambda \\
p &= 1 \\
A(\overrightarrow{\theta}) &= - \sum_{i=1}^{n} t_i \lambda \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(g) Uniform distribution: $X_i \sim \mathcal{U}(0, \theta)$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{\mathbf{1}[0 < x_i < \theta]}{\theta} \\
&= \exp(\sum_{i=1}^{n} \log \mathbf{1}[0 < x_i < \theta] - n \log \theta)
\end{aligned}
$$

Since $\sum_{i=1}^{n} \log \mathbf{1}[0 < x_i < \theta]$ does not factor into a function of only $\vec{x}$ and θ, this is not an exponential family.

(h) Exponential distribution (hazard parameterization): $X_i \sim \mathcal{E}(\lambda)$ with $\theta = \lambda$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \lambda \exp(-\lambda x_i) \\
&= \exp(-\lambda \sum_{i=1}^{n} x_i + n \log \lambda) \qquad\qquad h(\overrightarrow{x}) = 1 \\
\overrightarrow{T}(\overrightarrow{x}) &= \sum_{i=1}^{n} x_i \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= -\lambda \\
p &= 1 \\
A(\overrightarrow{\theta}) &= n \log \lambda \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(i) Exponential distribution (mean parameterization): $X_i \sim \mathcal{E}(\mu)$ with $\theta = \mu$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \lambda^{-1} \exp(-\lambda^{-1} x_i) \\
&= \exp(-\lambda^{-1} \sum_{i=1}^{n} x_i + -n \log \lambda) \qquad\qquad h(\overrightarrow{x}) = 1 \\
\overrightarrow{T}(\overrightarrow{x}) &= \sum_{i=1}^{n} x_i \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= -\lambda^{-1} \\
p &= 1 \\
A(\overrightarrow{\theta}) &= -n \log \lambda \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(j) Gamma distribution (unshifted): $X_i \sim \Gamma(\alpha, \beta, A = 0)$ with $\theta = (\alpha, \beta)$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{e^{-x_i/\beta} x_i^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} \\
&= \exp\left(\frac{-\sum_{i=1}^{n} x_i}{\beta} + (\alpha - 1)\sum_{i=1}^{n} \log x_i - n\log(\beta^{\alpha}\Gamma(\alpha)) \right) \\
&h(\overrightarrow{x}) = 1 \\
&\overrightarrow{T}(\overrightarrow{x}) = (\sum_{i=1}^{n} x_i,\ \sum_{i=1}^{n} \log x_i) \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = (-\frac{1}{\beta}, \alpha - 1) \\
&p = 2 \\
&A(\overrightarrow{\theta}) = -n\log(\beta^{\alpha}\Gamma(\alpha))
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved. $dim(\overrightarrow{T}(\overrightarrow{x})) = 2$.

(k) Normal distribution with known variance: $X_i \sim \mathcal{N}(\mu, \sigma^2)$ with $\theta = \mu$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \\
&h(\overrightarrow{x}) = e^{-\frac{\sum_{i=1}^{n} x_i^2}{2\sigma^2}} \\
&\overrightarrow{T}(\overrightarrow{x}) = \frac{\sum_{i=1}^{n} x_i}{\sigma^2} \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = \mu \\
&p = 1 \\
&A(\overrightarrow{\theta}) = -\frac{n\mu^2}{2\sigma^2} - n\log(\sqrt{2\pi}\sigma) \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 1.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(l) Normal distribution with unknown variance: $X_i \sim \mathcal{N}(\mu, \sigma^2)$ with $\overrightarrow{\theta} = (\mu, \sigma^2)$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \\
h(\overrightarrow{x}) &= 1 \\
\overrightarrow{T}(\overrightarrow{x}) &= (-\sum_{i=1}^{n}\frac{x_i^2}{2}, \sum_{i=1}^{n} x_i) \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= (\frac{1}{\sigma^2}, \frac{\mu}{\sigma}) \\
p &= 2 \\
A(\overrightarrow{\theta}) &= -\frac{n\mu^2}{2\sigma^2} - n\log(\sqrt{2\pi}\sigma) \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 2
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(m) Normal distribution with specified mean-variance relationship: $X_i \sim \mathcal{N}(\mu, \mu)$ with $\theta = \mu$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\mu}}}{\sqrt{2\pi}\sqrt{\mu}} \\
h(\overrightarrow{x}) &= e^{\sum_{i=1}^{n} x_i} \\
\overrightarrow{T}(\overrightarrow{x}) &= -\frac{\sum_{i=1}^{n} x_i^2}{2} \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= -\frac{1}{2\eta} \\
p &= 1 \\
A(\overrightarrow{\theta}) &= -\frac{n}{2}\mu - \frac{n}{2}\log(\sqrt{2\pi\mu}) \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 1
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(n) Normal distribution with specified mean-variance relationship: $X_i \sim \mathcal{N}(\mu, \mu^2)$ with $\theta = \mu$.

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\mu^2}}}{\sqrt{2\pi}\mu} \\
&h(\overrightarrow{x}) = 1 \\
&\overrightarrow{T}(\overrightarrow{x}) = (-\frac{\sum_{i=1}^{n} x_i^2}{2}, \sum_{i=1}^{n} x_i) \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = (\frac{1}{\mu^2}, \frac{1}{\mu}) \\
&p = 2 \\
&A(\overrightarrow{\theta}) = -\frac{n}{2} - n\log(\sqrt{2\pi}\mu) \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 2
\end{aligned}
$$

In this case, dimension of θ is 1 but $p = 2$. Therefore, this family is a curved exponential family.

(o) Normal distribution with systematically varying means: $X_i \sim \mathcal{N}(c_i\mu, \sigma^2)$ with $\overrightarrow{\theta} = (\mu, \sigma^2)$ (c_i-s are known).

Solution.

$$
\begin{aligned}
&f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-c_i\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \\
&h(\overrightarrow{x}) = 1 \\
&\overrightarrow{T}(\overrightarrow{x}) = (-\frac{\sum_{i=1}^{n} x_i^2}{2}, \sum_{i=1}^{n} c_i x_i) \\
&\overrightarrow{\eta}(\overrightarrow{\theta}) = (\frac{1}{\sigma^2}, \frac{\mu}{\sigma^2}) \\
&p = 2 \\
&A(\overrightarrow{\theta}) = -\frac{\sum_{i=1}^{n} c_i^2\mu^2}{2\sigma^2} - n\log(\sqrt{2\pi}\sigma) \\
&dim(\overrightarrow{T}(\overrightarrow{x})) = 2
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(p) Normal distribution with systematically varying variances: $X_i \sim \mathcal{N}(c_i\mu, \sigma^2)$ with $\overrightarrow{\theta} = (\mu, c_i\sigma^2)$ (c_i-s are known).

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2c_i\sigma^2}}}{\sqrt{2\pi c_i}\sigma} \\
h(\overrightarrow{x}) &= 1 \\
\overrightarrow{T}(\overrightarrow{x}) &= (-\sum_{i=1}^{n}\frac{x_i^2}{2c_i}, \sum_{i=1}^{n}\frac{x_i}{c_i}) \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= (\frac{1}{\sigma^2}, \frac{\mu}{\sigma^2}) \\
p &= 2 \\
A(\overrightarrow{\theta}) &= -\frac{1}{2}\sum_{i=1}^{n}\log c_i - n\log(\sqrt{2\pi}\sigma) - \frac{1}{2}\sum_{i=1}^{n}\frac{\mu^2}{c_i\sigma^2} \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 2.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(q) Lognormal distributions: $X_i \sim \mathcal{LN}(\mu, \sigma^2)$ with $\overrightarrow{\theta} = (\mu, \sigma^2)$.

Solution.

$$
\begin{aligned}
f_{\overrightarrow{X}}(\overrightarrow{x}) &= \prod_{i=1}^{n} \frac{1}{x_i\sqrt{2\pi}\sigma} e^{-\frac{(\log x-\mu)^2}{2\sigma^2}} \\
h(\overrightarrow{x}) &= \prod_{i=1}^{n}\frac{1}{x_i} \\
\overrightarrow{T}(\overrightarrow{x}) &= (\sum_{i=1}^{n}(\log x_i)^2, \sum_{i=1}^{n}(\log x_i)) \\
\overrightarrow{\eta}(\overrightarrow{\theta}) &= (-\frac{1}{2\sigma^2}, \frac{\mu}{\sigma^2}) \\
p &= 2 \\
A(\overrightarrow{\theta}) &= -\frac{n\mu^2}{2\sigma^2} - n\log(\sqrt{2\pi}\sigma) \\
dim(\overrightarrow{T}(\overrightarrow{x})) &= 2.
\end{aligned}
$$

Since dimension of $\theta = p$, this family is not curved.

(r) Weibull distribution: $X_i \sim Weib(p, \lambda)$ with $\overrightarrow{\theta} = (p, \lambda)$.

Solution.

$$
f_{\overrightarrow{X}}(\overrightarrow{x}) = \prod_{i=1}^{n} \frac{p}{\lambda}\left(\frac{x_i}{\lambda}\right)^{p-1} e^{-\left(\frac{x_i}{\lambda}\right)^p} I(x_i \geq 0)
$$

Since the term $-\left(\frac{x_i}{\lambda}\right)^p$ can not be written as a product of $T(x)$ and $\eta(\theta)$, (since p is an element of θ), this distribution does not belong to any exponential family.

MORE INVOLVED PROBLEMS

4. Let X be a continuous random variable with density $f_X(x|,\theta)$ for some real θ. Furthermore, suppose

   - $A = \{x : f_X(x|\theta) > 0\}$ (the "support" of the distribution of X) does not depend on θ, and
   - we can twice interchange the order of integration with respect to x and differentiation with respect to θ, so

$$\frac{\partial}{\partial\theta}\int f_X(x|\theta)dx = \int \left(\frac{\partial}{\partial\theta}f_X(x|\theta)\right)dx \tag{1}$$

$$\frac{\partial^2}{\partial\theta^2}\int f_X(x|\theta)dx = \int \left(\frac{\partial^2}{\partial\theta^2}f_X(x|\theta)\right)dx. \tag{2}$$

Consider the "efficient score" (transformation)

$$U(X) = \frac{\partial}{\partial\theta}\log f_X(X|\theta)$$

(a) Find the expectation

$$\mu_U = E[U(X)] = \int_{-\infty}^{\infty} U(x)f_X(x|\theta)dx.$$

(b) Show that

$$Var(U(X)) = \int_{-\infty}^{\infty}(U(x)-\mu_U)^2 f_X(x|\theta)dx = -E[\frac{\partial^2}{\partial\theta^2}\log f_X(X|\theta)]$$

(Note that this problem can be solved in the exact same way if X were a discrete random variable by substituting sums for integrals.)

Solution.

(a)

$$\begin{aligned}
&E[U(X)] \\
&= \int_{-\infty}^{\infty}\left(\frac{\partial}{\partial\theta}\log f_X(x|\theta)\right)f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty}\frac{\frac{\partial}{\partial\theta}f_X(x|\theta)}{f_X(x|\theta)}f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}f_X(x|\theta)dx \\
&= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f_X(x|\theta)dx \text{ (from (1))} \\
&= \frac{\partial}{\partial\theta}1 = 0.
\end{aligned}$$

(b) From part (a), we get that $\mu_U = 0$. Hence,

$$Var(U(X)) = \int_{-\infty}^{\infty} U(x)^2 f_X(x|\theta)dx \tag{3}$$

$$\begin{aligned}
&E[\frac{\partial^2}{\partial\theta^2}\log f_X(X|\theta)] \\
&= \int_{-\infty}^{\infty} \frac{\partial^2}{\partial\theta^2}\log(f_X(x|\theta))f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}\log(f_X(x|\theta))\right)f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial\theta}\left(\frac{\frac{\partial}{\partial\theta}f_X(x|\theta)}{f_X(x|\theta)}\right)f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty} \left[\frac{\frac{\partial^2}{\partial\theta^2}f_X(x|\theta)}{f_X(x|\theta)} - \frac{\partial}{\partial\theta}f_X(x|\theta)\frac{\frac{\partial}{\partial\theta}f_X(x|\theta)}{f_X(x|\theta)^2}\right]f_X(x|\theta)dx \\
&= \frac{\partial^2}{\partial\theta^2}\int_{-\infty}^{\infty} f_X(x|\theta)dx \text{ (Using (2))} \\
&- \int_{-\infty}^{\infty}\left[\frac{\frac{\partial}{\partial\theta}f_X(x|\theta)}{f_X(x|\theta)}\right]^2 f_X(x|\theta)dx \\
&= 0 - \int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\log f_X(x|\theta)\right]^2 f_X(x|\theta)dx \\
&= -\int_{-\infty}^{\infty} U(x)^2 f_X(x|\theta)dx \\
&= -Var(U(X))
\end{aligned}$$

using (3).
Note: Conditions (1) *and* (2) *are important. In general, the exchange of integral and differentiation may not be possible. 1 point will be deducted for not mentioning any of them.*

5. Let X be a normally distributed random variable $X \sim N(\mu, \sigma^2)$ with σknown and unknown $\theta = \mu$.

   (a) Derive the efficient score U (as given in the previous problem) for this distribution.

   (b) Now suppose that the efficient score derived from the normal distribution were used to transform some other random variable Y having mean μ and variance τ^2 What is are the mean and variance of $U(Y)$?

Solution.

(a) Here $\theta = \mu$. The score is

$$\begin{aligned}
U(x) &= \frac{\partial}{\partial\theta} \log \left[e^{-\frac{(x-\theta)^2}{2\sigma^2} - \log\sqrt{2\pi}\sigma} \right] \\
&= \frac{\partial}{\partial\theta} \left(-\frac{(x-\theta)^2}{2\sigma^2} - \log\sqrt{2\pi}\sigma \right) \\
&= \frac{x-\theta}{\sigma^2}.
\end{aligned}$$

(b) $U(Y) = \dfrac{Y-\theta}{\sigma^2}$.

$$EU(Y) = \frac{\mu-\theta}{\sigma^2} = 0$$

since $\theta = \mu$.

$$Var(U(Y)) = \frac{Var(Y)}{\sigma^4} = \frac{\tau^2}{\sigma^4}.$$