9. Let $X_1, \dots, X_n$ be a sample from a population with density

$$\begin{aligned} f_\theta(x) &= a(\theta)h(x) \text{ if } \theta_1 \le x \le \theta_2 \\ &= 0 \text{ otherwise} \end{aligned}$$

where $h(x) \ge 0$, $\theta = (\theta_1, \theta_2)$ with $-\infty < \theta_1 \le \theta_2 < \infty$, and $a(\theta) = \left[\int_{\theta_1}^{\theta_2} h(x)dx\right]^{-1}$ is assumed to exist. Find a two-dimensional sufficient statistic for this problem and apply your result to the $\mathcal{U}[\theta_1, \theta_2]$ family of distributions.

10. Suppose $X_1, \dots, X_n$ are i.i.d. with density $f(x, \theta) = \frac{1}{2}e^{-|x-\theta|}$. Show that $(X_{(1)}, \dots, X_{(n)})$, the order statistics, are minimal sufficient.

Hint: $\frac{\partial}{\partial\theta} L_{\mathbf{X}}(\theta) = -\sum_{i=1}^n \operatorname{sgn}(X_i - \theta)$, $\theta \notin \{X_1, \dots, X_n\}$, which determines $X_{(1)}, \dots, X_{(n)}$.

11. Let $X_1, X_2, \dots, X_n$ be a sample from the uniform, $\mathcal{U}(0, \theta)$, distribution. Show that $X_{(n)} = \max\{X_i;\ 1 \le i \le n\}$ is minimal sufficient for θ.

12. *Dynkin, Lehmann, Scheffé's Theorem.* Let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ where P_θ is discrete concentrated on $\mathcal{X} = \{x_1, x_2, \dots\}$. Let $p(x, \theta) \equiv P_\theta[X = x] \equiv L_x(\theta) > 0$ on $\mathcal{X}$. Show that $\frac{L_X(\cdot)}{L_X(\theta_0)}$ is minimial sufficient.

Hint: Apply the factorization theorem.

13. Suppose that $\mathbf{X} = (X_1, \dots, X_n)$ is a sample from a population with continuous distribution function $F(x)$. If $F(x)$ is $N(\mu, \sigma^2)$, $T(\mathbf{X}) = (\bar{X}, \widehat{\sigma}^2)$, where $\widehat{\sigma}^2 = n^{-1}\sum(X_i - \bar{X})^2$, is sufficient, and $S(\mathbf{X}) = (X'_{(1)}, \dots, X'_{(n)})$, where $X'_{(i)} = (X_{(i)} - \bar{X})/\widehat{\sigma}$, is "irrelevant" (ancillary) for (μ, σ^2). However, $S(\mathbf{X})$ is exactly what is needed to estimate the "shape" of $F(x)$ when $F(x)$ is unknown. The shape of F is represented by the equivalence class $\mathcal{F} = \{F((\cdot - a)/b) : b > 0,\ a \in R\}$. Thus a distribution G has the same shape as F iff $G \in \mathcal{F}$. For instance, one "estimator" of this shape is the scaled empirical distribution function

$$\begin{aligned} \widehat{F}_s(x) &= j/n,\ x'_{(j)} \le x < x'_{(j+1)},\ j = 1, \dots, n-1 \\ &= 0,\ x < x'_{(1)} \\ &= 1,\ x \ge x'_{(n)}. \end{aligned}$$

Show that for fixed x, $\widehat{F}_s((x - \bar{x})/\widehat{\sigma})$ converges in probability to $F(x)$. Here we are using F to represent $\mathcal{F}$ because every member of $\mathcal{F}$ can be obtained from F.

14. *Kolmogorov's Theorem.* We are given a regular model with Θ finite.

(a) Suppose that a statistic $T(\mathbf{X})$ has the property that for any prior distribution on $\boldsymbol{\theta}$, the posterior distribution of $\boldsymbol{\theta}$ depends on $\mathbf{x}$ only through $T(\mathbf{x})$. Show that $T(\mathbf{X})$ is sufficient.

(b) Conversely show that if $T(\mathbf{X})$ is sufficient, then, for any prior distribution, the posterior distribution depends on $\mathbf{x}$ only through $T(\mathbf{x})$.

Hint: Apply the factorization theorem.

15. Let $X_1, \dots, X_n$ be a sample from $f(x - \theta)$, $\theta \in R$. Show that the order statistics are minimal sufficient when f is the density *Cauchy* $f(t) = 1/\pi(1 + t^2)$.

16. Let $X_1, \dots, X_m; Y_1, \dots, Y_n$ be independently distributed according to $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{N}(\eta, \tau^2)$, respectively. Find minimal sufficient statistics for the following three cases:

(i) μ, η, σ, τ are arbitrary: $-\infty < \mu, \eta < \infty$, $0 < \sigma, \tau$.

(ii) $\sigma = \tau$ and μ, η, σ are arbitrary.

(iii) $\mu = \eta$ and μ, σ, τ are arbitrary.

17. In Example 1.5.4, express t_1 as a function of $L_x(0, 1)$ and $L_x(1, 1)$.

Problems to Section 1.6

1. Prove the assertions of Table 1.6.1.

2. Suppose $X_1, \dots, X_n$ is as in Problem 1.5.3. In each of the cases (a), (b) and (c), show that the distribution of $\mathbf{X}$ forms a one-parameter exponential family. Identify η, B, T, and h.

3. Let X be the number of failures before the first success in a sequence of Bernoulli trials with probability of success θ. Then $P_\theta[X = k] = (1 - \theta)^k \theta$, $k = 0, 1, 2, \dots$ This is called the *geometric distribution* ($\mathcal{G}(\theta)$).

(a) Show that the family of geometric distributions is a one-parameter exponential family with $T(x) = x$.

(b) Deduce from Theorem 1.6.1 that if $X_1, \dots, X_n$ is a sample from $\mathcal{G}(\theta)$, then the distributions of $\sum_{i=1}^n X_i$ form a one-parameter exponential family.

(c) Show that $\sum_{i=1}^n X_i$ in part (b) has a *negative binomial* distribution with parameters (n, θ) defined by $P_\theta[\sum_{i=1}^n X_i = k] = \binom{n + k - 1}{k} (1 - \theta)^k \theta^n$, $k = 0, 1, 2, \dots$ (The negative binomial distribution is that of the number of failures before the nth success in a sequence of Bernoulli trials with probability of success θ.)

Hint: By Theorem 1.6.1, $P_\theta[\sum_{i=1}^n X_i = k] = c_k (1 - \theta)^k \theta^n$, $0 < \theta < 1$. If

$$\sum_{k=0}^{\infty} c_k \omega^k = \frac{1}{(1 - \omega)^n}, \quad 0 < \omega < 1, \text{ then } c_k = \frac{1}{k!} \frac{d^k}{d\omega^k} (1 - \omega)^{-n} \bigg|_{\omega = 0}.$$

4. Which of the following families of distributions are exponential families? (Prove or disprove.)

(a) The $\mathcal{U}(0, \theta)$ family

(b) $p(x, \theta) = \{\exp[-2\log\theta + \log(2x)]\}1[x \in (0, \theta)]$

(c) $p(x, \theta) = \frac{1}{9}$, $x \in \{0.1 + \theta, \ldots, 0.9 + \theta\}$

(d) The $\mathcal{N}(\theta, \theta^2)$ family, $\theta > 0$

(e) $p(x, \theta) = 2(x + \theta)/(1 + 2\theta)$, $0 < x < 1$, $\theta > 0$

(f) $p(x, \theta)$ is the conditional frequency function of a binomial, $\mathcal{B}(n, \theta)$, variable X, given that $X > 0$.

5. Show that the following families of distributions are two-parameter exponential families and identify the functions η, B, T, and h.

(a) The beta family.

(b) The gamma family.

6. Let X have the Dirichlet distribution, $\mathcal{D}(\alpha)$, of Problem 1.2.15.

Show the distribution of X form an r-parameter exponential family and identify η, B, T, and h.

7. Let $\mathbf{X} = ((X_1, Y_1), \ldots, (X_n, Y_n))$ be a sample from a bivariate normal population. Show that the distributions of $\mathbf{X}$ form a five-parameter exponential family and identify η, B, T, and h.

8. Show that the family of distributions of Example 1.5.3 is not a one parameter exponential family.

Hint: If it were, there would be a set A such that $p(x, \theta) > 0$ on A for all θ.

9. Prove the analogue of Theorem 1.6.1 for discrete k-parameter exponential families.

10. Suppose that $f(x, \theta)$ is a positive density on the real line, which is continuous in x for each θ and such that if (X_1, X_2) is a sample of size 2 from $f(\cdot, \theta)$, then $X_1 + X_2$ is sufficient for θ. Show that $f(\cdot, \theta)$ corresponds to a one-arameter exponential family of distributions with $T(x) = x$.

Hint: There exist functions $g(t, \theta)$, $h(x_1, x_2)$ such that $\log f(x_1, \theta) + \log f(x_2, \theta) = g(x_1 + x_2, \theta) + h(x_1, x_2)$. Fix θ_0 and let $r(x, \theta) = \log f(x, \theta) - \log f(x, \theta_0)$, $q(x, \theta) = g(x, \theta) - g(x, \theta_0)$. Then, $q(x_1 + x_2, \theta) = r(x_1, \theta) + r(x_2, \theta)$, and hence, $[r(x_1, \theta) - r(0, \theta)] + [r(x_2, \theta) - r(0, \theta)] = r(x_1 + x_2, \theta) - r(0, \theta)$.

11. Use Theorems 1.6.2 and 1.6.3 to obtain moment-generating functions for the sufficient statistics when sampling from the following distributions.

(a) normal, $\boldsymbol{\theta} = (\mu, \sigma^2)$

(b) gamma, $\Gamma(p, \lambda)$, $\theta = \lambda$, p fixed

(c) binomial

(d) Poisson

(e) negative binomial (see Problem 1.6.3)

(f) gamma, $\Gamma(p, \lambda)$, $\boldsymbol{\theta} = (p, \lambda)$.