where $e_1, e_2, \ldots, e_n$ are iid with common pdf $f(x)$ and with support $(-\infty, \infty)$. Then the common pdf of X_i is $f_X(x;\theta) = f(x-\theta)$. We call model (6.2.7) a **location model**. Assume that $f(x)$ satisfies the regularity conditions. Then the information is

$$\begin{aligned} I(\theta) &= \int_{-\infty}^{\infty} \left(\frac{f'(x-\theta)}{f(x-\theta)}\right)^2 f(x-\theta)\,dx \\ &= \int_{-\infty}^{\infty} \left(\frac{f'(z)}{f(z)}\right)^2 f(z)\,dz, \end{aligned} \tag{6.2.8}$$

where the last equality follows from the transformation $z = x - \theta$. Hence, in the location model, the information does not depend on θ.

As an illustration, reconsider Example 6.1.1 concerning the Laplace distribution. Let $X_1, X_2, \ldots, X_n$ be a random sample from this distribution. Then it follows that X_i can be expressed as

$$X_i = \theta + e_i, \tag{6.2.9}$$

where $e_1, \ldots, e_n$ are iid with common pdf $f(z) = 2^{-1}\exp\{-|z|\}$, for $-\infty < z < \infty$. As we did in Example 6.1.1, use $\frac{d}{dz}|z| = \mathrm{sgn}(z)$. Then $f'(z) = -2^{-1}\mathrm{sgn}(z)\exp\{-|z|\}$ and, hence, $[f'(z)/f(z)]^2 = [-\mathrm{sgn}(z)]^2 = 1$, so that

$$I(\theta) = \int_{-\infty}^{\infty} \left(\frac{f'(z)}{f(z)}\right)^2 f(z)\,dz = \int_{-\infty}^{\infty} f(z)\,dz = 1. \tag{6.2.10}$$

Note that the Laplace pdf does not satisfy the regularity conditions, but this argument can be made rigorous; see Huber (1981) and also Chapter 10. ■

From (6.2.6), for a sample of size 1, say X_1, Fisher information is the variance of the random variable $\frac{\partial \log f(X_1;\theta)}{\partial\theta}$. What about a sample of size n? Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having pdf $f(x;\theta)$. The likelihood $L(\theta)$ is the pdf of the random sample, and the random variable whose variance is the information in the sample is given by

$$\frac{\partial \log L(\theta, \mathbf{X})}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial \log f(X_i;\theta)}{\partial\theta}.$$

The summands are iid with common variance $I(\theta)$. Hence the information in the sample is

$$\mathrm{Var}\left(\frac{\partial \log L(\theta, \mathbf{X})}{\partial\theta}\right) = nI(\theta). \tag{6.2.11}$$

Thus the information in a random sample of size n is n times the information in a sample of size 1. So, in Example 6.2.1, the Fisher information in a random sample of size n from a Bernoulli $b(1,\theta)$ distribution is $n/[\theta(1-\theta)]$.

We are now ready to obtain the Rao–Cramér lower bound, which we state as a theorem.

Theorem 6.2.1 (Rao–Cramér Lower Bound). *Let $X_1, \ldots, X_n$ be iid with common pdf $f(x;\theta)$ for $\theta \in \Omega$. Assume that the regularity conditions (R0)-(R4) hold. Let $Y = u(X_1, X_2, \ldots, X_n)$ be a statistic with mean $E(Y) = E[u(X_1, X_2, \ldots, X_n)] = k(\theta)$. Then*

$$Var(Y) \geq \frac{[k'(\theta)]^2}{nI(\theta)}. \tag{6.2.12}$$

Proof: The proof is for the continuous case, but the proof for the discrete case is quite similar. Write the mean of Y as

$$k(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \ldots, x_n) f(x_1;\theta)\cdots f(x_n;\theta)\,dx_1\cdots dx_n.$$

Differentiating with respect to θ, we obtain

$$\begin{aligned} k'(\theta) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) \left[\sum_1^n \frac{1}{f(x_i;\theta)} \frac{\partial f(x_i;\theta)}{\partial\theta}\right] \\ &\qquad \times f(x_1;\theta)\cdots f(x_n;\theta)\,dx_1\cdots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) \left[\sum_1^n \frac{\partial \log f(x_i;\theta)}{\partial\theta}\right] \\ &\qquad \times f(x_1;\theta)\cdots f(x_n;\theta)\,dx_1\cdots dx_n. \end{aligned} \tag{6.2.13}$$

Define the random variable Z by $Z = \sum_1^n [\partial \log f(X_i;\theta)/\partial\theta]$. We know from (6.2.2) and (6.2.11) that $E(Z) = 0$ and $\mathrm{Var}(Z) = nI(\theta)$, respectively. Also, equation (6.2.13) can be expressed in terms of expectation as $k'(\theta) = E(YZ)$. Hence we have

$$k'(\theta) = E(YZ) = E(Y)E(Z) + \rho\sigma_Y\sqrt{nI(\theta)},$$

where ρ is the correlation coefficient between Y and Z. Using $E(Z) = 0$, this simplifies to

$$\rho = \frac{k'(\theta)}{\sigma_Y\sqrt{nI(\theta)}}.$$

Because $\rho^2 \leq 1$, we have

$$\frac{[k'(\theta)]^2}{\sigma_Y^2 nI(\theta)} \leq 1,$$

which, upon rearrangement, is the desired result. ■

Corollary 6.2.1. *Under the assumptions of Theorem 6.2.1, if $Y = u(X_1, \ldots, X_n)$ is an unbiased estimator of θ, so that $k(\theta) = \theta$, then the Rao-Cramér inequality becomes*

$$Var(Y) \geq \frac{1}{nI(\theta)}.$$