Theorem. Let X and Y be independently distributed random observables with density $f_1(x|\theta)$ and $f_2(y|\theta)$. If $I(\theta)$, $I_1(\theta)$ and $I_2(\theta)$ are the information numbers about θ contained in (X, Y), X and Y, respectively, then

$$I(\theta) = I_1(\theta) + I_2(\theta).$$

Proof. To see this additive property, note that,

$$\begin{aligned}
I(\theta) &= Var_\theta\left[\frac{\partial}{\partial\theta}\log f(X, Y|\theta)\right] = Var_\theta\left[\frac{\partial}{\partial\theta}\log\{f_1(X|\theta)f_2(Y|\theta)\}\right] \\
&= Var_\theta\left[\frac{\partial}{\partial\theta}\log f_1(X|\theta) + \frac{\partial}{\partial\theta}\log f_2(Y|\theta)\right] \\
&= Var_\theta\left[\frac{\partial}{\partial\theta}\log f_1(X|\theta)\right] + Var_\theta\left[\frac{\partial}{\partial\theta}\log f_2(Y|\theta)\right] \\
&= I_1(\theta) + I_2(\theta).
\end{aligned}$$

Example. Let $X_1, \ldots, X_n$ be i.i.d Bernoulli(p). Then the information in the sample is $I(p) = \frac{n}{p(1-p)}$ since the information in each of the observations is $I_1(p) = \frac{1}{p(1-p)}$. Further, $\sum_{i=1}^n X_i \sim$ Binomial(n, p) is sufficient for p and has the same likelihood function as that of the sample. Thus Binomial(n, p) has the same $I(p)$ of $\frac{n}{p(1-p)}$.

For multi-parameter problems, one defines the Information Matrix,

$$I(\theta) = ((I_{ij}(\theta))), \text{ where } I_{ij}(\theta) = E\left[\frac{\partial}{\partial\theta_i}\log f(X|\theta)\frac{\partial}{\partial\theta_j}\log f(X|\theta)\right].$$

$I(\theta)$ depends on the particular parametrization chosen: Suppose $\eta = c(\theta)$, where $c(.)$ is one-one and differentaible. Then $\theta = h(\eta) = c^{-1}(\eta)$. Therefore, letting $f^*(x|\eta) = f(x|\theta)|_{\theta=h(\eta)}$,

$$\begin{aligned}
I^*(\eta) &= E_\eta\left[\frac{\partial}{\partial\eta}\log f^*(X|\eta)\right]^2 \\
&= E_\eta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\,|_{\theta=h(\eta)}\,\frac{\partial}{\partial\eta}h(\eta)\right]^2 \\
&= E_\eta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]^2|_{\theta=h(\eta)}\left(\frac{\partial}{\partial\eta}h(\eta)\right)^2 \\
&= I(\theta)\left(\frac{d\theta}{d\eta}\right)^2|_{\theta=h(\eta)}.
\end{aligned}$$

Example. Let $f(x|\alpha) = \alpha x \exp\left(-\frac{\alpha}{2}x^2\right)$, $x > 0$, $\alpha > 0$. What is $I(\alpha)$? Using the definition, since

$$\begin{aligned} \log f(x|\alpha) &= \log \alpha + \log x - \frac{\alpha}{2}x^2, \\ \frac{\partial}{\partial \alpha} \log f(x|\alpha) &= \frac{1}{\alpha} - \frac{x^2}{2}, \text{ so} \\ I(\alpha) &= E_\alpha \left[\left(\frac{X^2}{2} - \frac{1}{\alpha} \right)^2 \right], \end{aligned}$$

which is difficult to compute, but using the alternative formula, we get,

$$\frac{\partial^2}{\partial \alpha^2} \log f(x|\alpha) = -\frac{1}{\alpha^2},$$

so that $I(\alpha) = 1/\alpha^2$. Sometimes, reparametrization can help too.

Example. Let $f(x|\theta) = \frac{x}{\theta^2} \exp\left(-\frac{x^2}{2\theta^2}\right)$, $x > 0$, $\theta > 0$. Find $I(\theta)$. We note

$$\begin{aligned} \log f(x|\theta) &= \log x - 2\log\theta - \frac{x^2}{2\theta^2}, \\ \frac{\partial}{\partial \theta} \log f(x|\theta) &= -\frac{2}{\theta} + \frac{x^2}{\theta^3} = \frac{1}{\theta^3}(x^2 - 2\theta^2), \text{ so} \\ I(\theta) &= \frac{1}{\theta^6} E_\theta \left(X^2 - 2\theta^2 \right)^2, \end{aligned}$$

which is again difficult to compute. We may look at

$$\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) = \frac{2}{\theta^2} - \frac{3x^2}{\theta^4} = -\frac{1}{\theta^4}(3x^2 - 2\theta^2),$$

and try to compute $I(\theta) = \frac{1}{\theta^4} E_\theta(3X^3 - 2\theta^2)$, which doesn't seem to simplify the calculation. Instead, let $\alpha = \alpha(\theta) = 1/\theta^2$. Then from the previous example, $I(\alpha) = 1/\alpha^2$. Therefore,

$$I^*(\theta) = I(\alpha(\theta))(\alpha'(\theta))^2 = \theta^4 \frac{4}{\theta^6} = \frac{4}{\theta^2}.$$

Example. Let $X \sim N(\mu, \sigma^2)$. Then $I(\mu, \sigma^2) = ((I_{ij}(\mu, \sigma^2)))$, where

$$\begin{aligned} I_{11}(\mu, \sigma^2) &= E_{\mu,\sigma^2} \left[\frac{\partial}{\partial \mu} \log f(X|\mu, \sigma^2) \right]^2 \\ I_{22}(\mu, \sigma^2) &= E_{\mu,\sigma^2} \left[\frac{\partial}{\partial \sigma^2} \log f(X|\mu, \sigma^2) \right]^2 \\ I_{12}(\mu, \sigma^2) &= E_{\mu,\sigma^2} \left[\frac{\partial}{\partial \mu} \log f(X|\mu, \sigma^2) \frac{\partial}{\partial \sigma^2} \log f(X|\mu, \sigma^2) \right]. \end{aligned}$$

Since

$$\begin{aligned}
\log f(x|\mu,\sigma^2) &= -\frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2, \\
\frac{\partial}{\partial\mu}\log f(x|\mu,\sigma^2) &= -\frac{1}{2\sigma^2}2(x-\mu)(-1), \\
\frac{\partial}{\partial\sigma^2}\log f(x|\mu,\sigma^2) &= -\frac{1}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2,
\end{aligned}$$

we obtain

$$\begin{aligned}
I_{11}(\mu,\sigma^2) &= E_{\mu,\sigma^2}\left[\frac{(X-\mu)^2}{\sigma^4}\right] = \frac{1}{\sigma^2}, \\
I_{22}(\mu,\sigma^2) &= \frac{1}{4\sigma^8}E_{\mu,\sigma^2}\left[(X-\mu)^2-\sigma^2\right]^2 = \frac{2\sigma^4}{4\sigma^8} = \frac{1}{2\sigma^4}. \\
I_{12}(\mu,\sigma^2) &= \frac{1}{2}E_{\mu,\sigma^2}\left[\left(\frac{X-\mu}{\sigma^2}\right)\left(\frac{(X-\mu)^2-\sigma^2}{\sigma^4}\right)\right] = 0.
\end{aligned}$$

Thus,

$$I(\mu,\sigma^2) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}.$$