556: MATHEMATICAL STATISTICS I

SCORE FUNCTION AND FISHER INFORMATION FOR LOCATION-SCALE FAMILIES

The location-scale family for rv X is defined using a linear transformation of a standard variable Z by

$$X = \mu + \sigma Z$$

for $\mu \in \mathbb{R}$ and $\sigma > 0$, and $f_Z(.)$ is a "standard" distribution that does not depend on any parameters.

- For a continuous rv, we have for the pdf

$$f_X(x;\mu,\sigma) = \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right)$$

- For a discrete rv, we have for the pmf

$$f_X(x;\mu,\sigma) = f_Z\left(\frac{x-\mu}{\sigma}\right).$$

  In most settings, the discrete location-scale family is not that useful as it effectively amounts merely to a re-labelling of the support of f_Z.

- In both cases, for the cdf, we have

$$F_X(x;\mu,\sigma) = F_Z\left(\frac{x-\mu}{\sigma}\right).$$

The score function, $\mathbf{S}(x;\theta)$, is defined by

$$\mathbf{S}(x;\theta) = \frac{\partial}{\partial\theta}\left\{\log f_X(x;\theta)\right\}.$$

If θ is m-dimensional, then $\mathbf{S}(X;\theta)$ is $(m \times 1)$. For the location-scale family, we have that $m = 2$.

For the continuous case, we consider the construction where the pdf f_Z has support $\mathbb{Z} = (a, b)$ for values $-\infty \leq a < b \leq \infty$. We have

$$\log f_X(x;\theta) \equiv \log f_X(x;\mu,\sigma) = -\log\sigma + \log f_Z\left(\frac{x-\mu}{\sigma}\right)$$

and so

$$S_\mu(x;\mu,\sigma) = \frac{\partial}{\partial\mu}\left\{\log f_X(x;\mu,\sigma)\right\} = \frac{\partial}{\partial\mu}\left\{\log f_Z((x-\mu)/\sigma)\right\} = -\frac{1}{\sigma}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} \tag{1}$$

$$S_\sigma(x;\mu,\sigma) = \frac{\partial}{\partial\sigma}\left\{\log f_X(x;\mu,\sigma)\right\} = -\frac{1}{\sigma} + \frac{\partial}{\partial\sigma}\left\{\log f_Z((x-\mu)/\sigma)\right\} = -\frac{1}{\sigma} - \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} \tag{2}$$

where

$$\dot{f}_Z(z) = \frac{\partial f_Z(z)}{\partial z}.$$

In the following calculations, integration is over the support (a, b). For the first score function (1): we have that

$$\begin{aligned}
\mathbb{E}_X\left[\frac{\dot{f}_Z((X-\mu)/\sigma)}{f_Z((X-\mu)/\sigma)}\right] &= \int \frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} f_X(x;\mu,\sigma)\,dx \\
&= \int \frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} \frac{1}{\sigma} f_Z((x-\mu)/\sigma)\,dx \\
&= \int \dot{f}_Z(z)\,dz \qquad\qquad z=(x-\mu)/\sigma \\
&= f_Z(b) - f_Z(a)
\end{aligned}$$

by standard calculus arguments. Note that this equates to zero if

$$\lim_{z\longrightarrow a} f_Z(z) = \lim_{z\longrightarrow b} f_Z(z) = 0.$$

which certainly holds if the support is the whole of $\mathbb{R}$.

For the second score function (2): we have that

$$\begin{aligned}
\mathbb{E}_X\left[(X-\mu)\frac{\dot{f}_Z((X-\mu)/\sigma)}{f_Z((X-\mu)/\sigma)}\right] &= \int (x-\mu)\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} f_X(x;\mu,\sigma)\,dx \\
&= \int (x-\mu)\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} \frac{1}{\sigma} f_Z((x-\mu)/\sigma)\,dx \\
&= \sigma \int z\dot{f}_Z(z)\,dz \qquad\qquad z=(x-\mu)/\sigma
\end{aligned}$$

and, using integration by parts

$$\int_a^b z\dot{f}_Z(z)\,dz = [zf_Z(z)]_a^b - \int_a^b f_Z(z)\,dz = (bf_Z(b) - af_Z(a)) - 1.$$

Note that if $a = -\infty$ and $b = \infty$, this calculation is still valid as

$$\lim_{z\longrightarrow -\infty} zf_Z(z) = \lim_{z\longrightarrow \infty} zf_Z(z) = 0$$

because $f_Z(z)$ is integrable, and therefore is $\text{o}(|z|^{-(1+\delta)})$ for $\delta > 0$ as $|z| \longrightarrow \infty$. Therefore, from (1) and (2), we have under the usual regularity conditions

$$\mathbb{E}_X\left[\mathbf{S}(X;\mu,\sigma)\right] = -\frac{1}{\sigma}\begin{bmatrix} f_Z(b) - f_Z(a) \\ (bf_Z(b) - af_Z(a)) \end{bmatrix}. \tag{3}$$

Notice that this reduces to zero whenever

$$f_Z(a) = f_Z(b) = 0$$

which is a common case when considering location-scale models (eg Normal, Cauchy etc).

The Fisher information, $\mathcal{I}(\theta)$, is then defined as

$$\mathcal{I}(\theta) = \text{Var}_X[\mathbf{S}(X;\theta)] = \mathbb{E}_X\left[\mathbf{S}(X;\theta)\mathbf{S}(X;\theta)^\top\right] - \mathbb{E}_X\left[\mathbf{S}(X;\theta)\right]\mathbb{E}_X\left[\mathbf{S}(X;\theta)\right]^\top \tag{4}$$

which is an $(m \times m)$ symmetric and non-negative definite matrix. The three distinct elements of $\mathbf{S}(x;\theta)\mathbf{S}(x;\theta)^\top$ are

$$\{S_\mu(x;\mu,\sigma)\}^2 = \frac{1}{\sigma^2}\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 \tag{5}$$

$$S_\mu(x;\mu,\sigma)S_\sigma(x;\mu,\sigma) = \left\{\frac{1}{\sigma}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}\left\{\frac{1}{\sigma} + \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\} \tag{6}$$

$$\{S_\sigma(x;\mu,\sigma)\}^2 = \left\{\frac{1}{\sigma} + \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 \tag{7}$$

From (5):

$$\begin{aligned}\mathbb{E}_X\left[\{S_\mu(X;\mu,\sigma)\}^2\right] &= \frac{1}{\sigma^2}\int\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 f_X(x;\mu,\sigma)\,dx \\ &= \frac{1}{\sigma^2}\int\frac{\{\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz \qquad z=(x-\mu)/\sigma.\end{aligned} \tag{8}$$

From (6):

$$\begin{aligned}\mathbb{E}_X\left[S_\mu(x;\mu,\sigma)S_\sigma(x;\mu,\sigma)\right] &= \int\left\{\frac{1}{\sigma}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}\left\{\frac{1}{\sigma} + \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\} f_X(x;\mu,\sigma)\,dx \\ &= \frac{1}{\sigma^2}\int\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}f_X(x;\mu,\sigma)\,dx \\ &\qquad + \frac{1}{\sigma^3}\int(x-\mu)\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 f_X(x;\mu,\sigma)\,dx \\ &= \frac{1}{\sigma^2}\int\dot{f}_Z(z)\,dz + \frac{1}{\sigma^2}\int z\frac{\{\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz \qquad z=(x-\mu)/\sigma \\ &= \frac{1}{\sigma^2}\int\frac{\dot{f}_Z(z)(z\dot{f}_Z(z)+f_Z(z))}{f_Z(z)}\,dz\end{aligned} \tag{9}$$

From (7):

$$\begin{aligned}\mathbb{E}_X\left[\{S_\sigma(X;\mu,\sigma)\}^2\right] &= \int\left\{\frac{1}{\sigma} + \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 f_X(x;\mu,\sigma)\,dx \\ &= \int\left\{\frac{1}{\sigma} + \frac{z}{\sigma}\frac{\dot{f}_Z(z)}{f_Z(z)}\right\}^2 f_Z(z)\,dz \qquad z=(x-\mu)/\sigma \\ &= \frac{1}{\sigma^2}\int\frac{\{f_Z(z)+z\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz.\end{aligned} \tag{10}$$

Therefore combining (3), (8), (9) and (10) we observe from the definition (4) that

$$\mathcal{I}(\theta) \equiv \mathcal{I}(\mu,\sigma) = \frac{1}{\sigma^2}\mathbf{V}_Z$$

where $\mathbf{V}_Z$ is a **constant** matrix computed from f_Z, with knowledge of the support (a,b).