MATH 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 may depend on other (fixed) parameters. We restrict attention to the continuous case: we have for the pdf

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

The score function, $\mathbf{S}(x;\theta)$, and Fisher information, $\mathcal{I}(\theta)$, are defined for θ $(m \times 1)$ by

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

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

For the location-scale family, we have that $m = 2$. Under standard regularity conditions we have that

$$\mathbb{E}_X[\mathbf{S}(X;\theta)] = \mathbf{0} \qquad \mathcal{I}(\theta) = -\mathbb{E}_X\left[\frac{\partial^2 \log f_X(X;\theta)}{\partial\theta\partial\theta^\top}\right] = -\mathbb{E}_X[\mathbf{\Psi}(X;\theta)].$$

We consider the case where the pdf f_Z has support $\mathbb{Z} = (a,b)$ for values $-\infty \le a < b \le \infty$. 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

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

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

where $\dot{f}_Z(z) = df_Z(z)/dz$. For the first score function (1):

$$\mathbb{E}_X\left[\frac{\dot{f}_Z((X-\mu)/\sigma)}{f_Z((X-\mu)/\sigma)}\right] = \int_{\mu+\sigma a}^{\mu+\sigma b} \frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} f_X(x;\mu,\sigma)\,dx = \int_{\mu+\sigma a}^{\mu+\sigma b} \frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\frac{1}{\sigma} f_Z((x-\mu)/\sigma)\,dx$$

$$= \int_a^b \dot{f}_Z(z)\,dz \qquad \text{setting } z = (x-\mu)/\sigma$$

$$= f_Z(b) - f_Z(a)$$

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

$$\mathbb{E}_X\left[(X-\mu)\frac{\dot{f}_Z((X-\mu)/\sigma)}{f_Z((X-\mu)/\sigma)}\right] = \int_{\mu+\sigma a}^{\mu+\sigma b}(x-\mu)\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}f_X(x;\mu,\sigma)\,dx$$

$$= \int_{\mu+\sigma a}^{\mu+\sigma b}(x-\mu)\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\frac{1}{\sigma}f_Z((x-\mu)/\sigma)\,dx = \sigma\int_a^b z\dot{f}_Z(z)\,dz$$

setting $z=(x-\mu)/\sigma$. 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 $\mathrm{o}(|z|^{-(1+\delta)})$ for $\delta>0$ as $|z|\longrightarrow\infty$. Therefore, from (1) and (2), we have under the usual regularity conditions

$$\mathbb{E}_X\left[\mathbf{S}(X;\mu,\sigma)\right] = -\frac{1}{\sigma}\begin{bmatrix} f_Z(b)-f_Z(a) \\ (bf_Z(b)-af_Z(a)) \end{bmatrix} = \mathbf{0} \tag{3}$$

whenever $f_Z(a)=f_Z(b)=0$ which often holds for location-scale models (eg Normal, Cauchy etc).

For the Fisher information, the three distinct elements of $\mathbf{S}(x;\theta)\mathbf{S}(x;\theta)^\top$ are

$$\{S_\mu(x;\mu,\sigma)\}^2 = \frac{1}{\sigma^2}\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 \tag{4}$$

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

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

- From (4):

$$\mathbb{E}_X\left[\{S_\mu(X;\mu,\sigma)\}^2\right] = \frac{1}{\sigma^2}\int_{\mu+\sigma a}^{\mu+\sigma b}\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 f_X(x;\mu,\sigma)\,dx = \frac{1}{\sigma^2}\int_a^b\frac{\{\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz \tag{7}$$

- From (5):

$\mathbb{E}_X\left[S_\mu(x;\mu,\sigma)S_\sigma(x;\mu,\sigma)\right]$

$$= \int_{\mu+\sigma a}^{\mu+\sigma b}\left\{\frac{1}{\sigma}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}\left\{\frac{1}{\sigma}+\frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}f_X(x;\mu,\sigma)\,dx$$

$$= \frac{1}{\sigma^2}\int_{\mu+\sigma a}^{\mu+\sigma b}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}f_X(x;\mu,\sigma)\,dx + \frac{1}{\sigma^3}\int_{\mu+\sigma a}^{\mu+\sigma b}(x-\mu)\left\{\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)}\right\}^2 f_X(x;\mu,\sigma)\,dx$$

$$= \frac{1}{\sigma^2}\int_a^b\dot{f}_Z(z)\,dz + \frac{1}{\sigma^2}\int_a^b z\frac{\{\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz$$

$$= \frac{1}{\sigma^2}\int_a^b\frac{\dot{f}_Z(z)(z\dot{f}_Z(z)+f_Z(z))}{f_Z(z)}\,dz \tag{8}$$

- From (6):

$$\mathbb{E}_X\left[\{S_\sigma(X;\mu,\sigma)\}^2\right] = \int_{\mu+\sigma a}^{b} \left\{ \frac{1}{\sigma} + \frac{(x-\mu)}{\sigma^2}\frac{\dot{f}_Z((x-\mu)/\sigma)}{f_Z((x-\mu)/\sigma)} \right\}^2 f_X(x;\mu,\sigma)\,dx$$

$$= \int_a^b \left\{ \frac{1}{\sigma} + \frac{z}{\sigma}\frac{\dot{f}_Z(z)}{f_Z(z)} \right\}^2 f_Z(z)\,dz$$

$$= \frac{1}{\sigma^2}\int_a^b \frac{\{f_Z(z) + z\dot{f}_Z(z)\}^2}{f_Z(z)}\,dz. \tag{9}$$

Therefore combining (3), (7), (8) and (9) we observe 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).