Depth notes

April 7, 2014

1 Depth and missspecification of scale parameter

Student depth as per Mizera/ Muller:

$$d(\mu, \sigma) = \inf_{u \neq 0} P(u_1(y-u) + u_2((y-u)^2 - \sigma^2) \geq 0)$$

P being any prob. distribution.

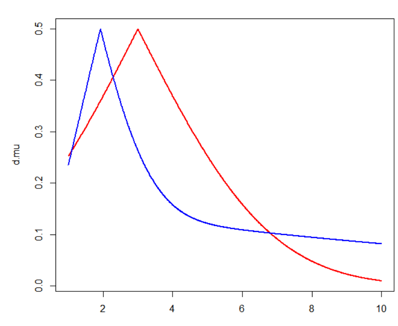
$N(\mu, \mu^2)$ example Curved normal location-scale family: $Y \sim N(\mu, \mu^2)$. Obtained by differentiating -ve of log-likelihood:

$$\begin{aligned} -l(y,\mu) &= \frac{(y-\mu)^2}{2\mu^2} + \log\mu + K = \frac{y^2}{2\mu^2} + \frac{y}{\mu} + \log\mu + K_1 \\ \Rightarrow -\frac{dl(y,\mu)}{d\mu} &= -\frac{y^2}{\mu^3} + \frac{y}{\mu^2} + \frac{1}{\mu} = \frac{1}{\mu^3}(\mu^2 + y\mu - y^2) \end{aligned}$$

Hence the Student depth at μ

$$\begin{aligned} d(\mu) &= \inf_{u \neq 0} P\left(\frac{u}{\mu^3}(\mu^2 + y\mu - y^2) \geq 0\right) \\ &= \min\left\{P\left(\frac{1}{\mu^3}(\mu^2 + y\mu - y^2) \geq 0\right), P\left(\frac{1}{\mu^3}(\mu^2 + y\mu - y^2) \leq 0\right)\right\} \\ &= \min\left\{P\left(\mu^2 + y\mu - y^2 \geq 0\right), P\left(\mu^2 + y\mu - y^2 \leq 0\right)\right\} \\ &= \min\left\{P\left[y \in \mu(1 \pm \sqrt{5})/2\right], 1 - P\left[y \in \mu(1 \pm \sqrt{5})/2\right]\right\} \end{aligned}$$

The plot below shows student depths for $\mu \in [1, 10]$ under the dependent model (blue, max at 1.93) and location model (red, max at 3). P is taken as $N(3, 3^2)$.



$N(\mu, f(\mu))$ **case** Assume now $f(\mu) > 0 \quad \forall \quad \mu \in \mathbb{R}$ differentiable. Then we have

$$\begin{aligned}
-l(\mu, y) &= \frac{(y-\mu)^2}{2f(\mu)} + \frac{\log f(\mu)}{2} + K \\
\Rightarrow -\frac{dl(y,\mu)}{d\mu} &= -\frac{(y-\mu)^2 f'(\mu)}{2f^2(\mu)} - \frac{y-\mu}{f(\mu)} + \frac{f'(\mu)}{2f(\mu)} \\
&= \frac{-y^2 + 2y\mu - \mu^2 - 2yf(\mu) + 2\mu f(\mu) + f(\mu)f'(\mu)}{2f^2(\mu)} \\
&= \frac{-y^2 f'(\mu) + 2y(\mu f'(\mu) - f(\mu)) - [\mu^2 f'(\mu) - 2\mu f(\mu) - f(\mu)f'(\mu)]}{2f^2(\mu)} \\
&= -\frac{(y-r_1)(y-r_2)}{2f^2(\mu)}
\end{aligned}$$

with r_1, r_2 being the two roots of the quadratic form in numerator:

$$r_1, r_2 = \frac{\mu f'(\mu) - f(\mu) \pm \sqrt{(\mu f'(\mu) - f(\mu))^2 - \mu^2 f'(\mu) + 2\mu f(\mu) + f(\mu) f'(\mu)}}{f'(\mu)} = ...$$

Proceeding similarly as the first case, we can conclude that the depth at μ will be

$$d(\mu) = \min\left\{P\left[y \in (r_{(1)}, r_{(2)})\right], 1 - P\left[y \in (r_{(1)}, r_{(2)})\right]\right\}$$

General location-scale family Same assumption $\sigma = k(\mu)$ and density $\exp(g(.))$. Then we have the negative log-likelihood:

$$\begin{aligned}
-l(\mu, y) &= -g\left(\frac{y-\mu}{k(\mu)}\right) + \log k(\mu) \\
\Rightarrow \quad -\frac{dl(y,\mu)}{d\mu} &= \frac{y-\mu}{k^2(\mu)} + \frac{1}{k(\mu)}\left[g'\left(\frac{y-\mu}{k(\mu)}\right) + k'(\mu)\right]
\end{aligned}$$

And the student depth comes out to be

$$\begin{aligned}
d(\mu) = \min\{&P[y - \mu + k(\mu)\tau(\mu; y) + k(\mu)k'(\mu) \geq 0], \\
&P[y - \mu + k(\mu)\tau(\mu; y) + k(\mu)k'(\mu) \leq 0]\}
\end{aligned}$$

with $\tau(\mu; y) = g((y-\mu)/k(\mu))$.

2 Empirical Likelihood depth

Tangent depth as per Mizera:

$$d(\theta) = \frac{1}{n} \min_{u \neq 0} \{\#i : u^T l_i'(\theta) \geq 0\}$$

with $l_i(\theta)$ being the log-likelihood corresponding to the i^{th} observation. In place of a parametric likelihood, we consider the empirical log-likelihood in the lines of Qin and Lawless, so that information about P_θ is obtained by r unbiased estimating eqns $g(X, \theta) = (g_1(X, \theta), ..., g_r(X, \theta))'$, with $Eg(X, \theta) = 0$. Thus we now have (See Qin, Lawless) for $i = 1, ..., n$

$$l_i(\theta) = \log\left[1 + t^T g(x_i, \theta)\right]$$

$t = t(\theta) \in \mathbb{R}^d$ being the solution of the equation

$$\sum_{i=1}^{n} \frac{g(x_i, \theta)}{1 + t^T g(x_i, \theta)} = 0_{d \times 1} \tag{1}$$

Hence we define the **Empirical likelihood depth** as

$$d_e(\theta) = \frac{1}{n} \min_{u \neq 0} \left\{ \#i : u^T \left(\frac{G(x_i, \theta) t}{1 + t^T g(x_i, \theta)} \right) \geq 0 \right\}$$

with t satisfying (1) and $G(x, \theta) \in \mathbb{R}^{p \times r}$ being the derivative matrix of $g^T(x, \theta)$ wrt θ.

$p = 1$ **case** For a single parameter θ, we shall have

$$d_e(\mu) = \frac{1}{n} \min \left\{ \left(\#i : \frac{G(x_i, \theta) t}{1 + t^T g(x_i, \theta)} \geq 0 \right), \left(\#i : \frac{G(x_i, \theta) t}{1 + t^T g(x_i, \theta)} \leq 0 \right) \right\}$$

Maximum likelihood depth Define the **Maximum empirical depth estimator** (MEDE) as the value of θ which maximizes the empirical likelihood depth evaluated there:

$$\hat{d}_e(\theta) = \arg\max_\theta d_e(\theta) = \arg\max_\theta \inf_{u \neq 0} P_n \left[u^T \left(\frac{G(x, \theta) t}{1 + t^T g(x, \theta)} \right) \geq 0 \right]$$

subject to the constraint (1).

Single parameter: univariate location problem $X_1, ..., X_n \sim F(x - \mu)$ iid; $\mu \in \mathbb{R}, g(X, \theta) = X - \mu$. Then $l_i(\mu) = \log[1 + t(x_i - \mu)]$, subject to (1). Hence

$$d_e(\mu) = \frac{1}{n} \min \left\{ \left(\#i : \frac{t}{1 + t(x_i - \mu)} \geq 0 \right), \left(\#i : \frac{t}{1 + t(x_i - \mu)} \leq 0 \right) \right\}$$

For a given μ we first solve for t from (1), then obtain the 'depth'. Solving for t is equivalent to minimizing $-\sum_i \log[1 + tg(x,\mu)]$. That can be done by Newton's method close to the actual mean μ_0, but if μ is far away it's minimized at ∞.

Single parameter: multiple estimating equations Normal example: $N(\mu, 1)$ population. We only know first moment μ, third moment $\mu^3 + 3\mu$. $g(X, \theta) = (X - \mu, X^3 - \mu^3 - 3\mu)^T$.

$$d_e(\mu) = \frac{1}{n} \min \left\{ \left(\#i : \frac{t_1 + (3\mu^2 + 3\mu) t_2}{1 + t^T g(x_i, \mu)} \geq 0 \right), \left(\#i : \frac{t_1 + (3\mu^2 + 3\mu) t_2}{1 + t^T g(x_i, \mu)} \leq 0 \right) \right\}$$