CASELLA AND BERGER NOTES AND PROBLEMS - CHAPTER 7

AUSTIN DAVID BROWN

1. Problems

Chapter 7: 1, 2, 6, 7, 8, 11, 12, 13, 14, 19, 20, 23, 24, 26, 33, 37, 38, 39, 41, 42, 44, 46
Chapter 7: 49, 50, 52(a,b), 57, 58, 59, 60

2. Notes

8 / 11 ____

8 / 12 ____

I think this is a good problem to work out.
Find the MLE for μ, σ^2 for a normal distribution.
Let $X = (X_k)_{k=1}^n$ be iid random vector with $X_k \sim N(\mu, \sigma^2)$.
Since $Int\mathbb{R} = \mathbb{R}$, the max, min must won't be on the boundary and the derivative will be 0.
Let $v = (\mu, \sigma^2)$.
The likelihood map $L_X : \mathbb{R}^2 \to \mathbb{R}$ is C^∞ and

$$L_X(v) = \frac{1}{(\pi 2\sigma^2)^{n/2}} \exp(\frac{-1}{2\sigma^2} \sum_{k=1}^n (x_k - \mu)^2).$$

Then

$$\log L_X(v) = \frac{-n}{2} \log(\pi 2\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - \mu)^2,$$

and

$$\begin{aligned} D(\log L_X(v))(v) &= \begin{pmatrix} \partial_\mu \log L_X(v) & \partial_{\sigma^2} \log L_X(v) \end{pmatrix} \\ &= \begin{pmatrix} \frac{n}{\sigma^2}(\overline{X_n} - \mu) & \frac{n}{2(\sigma^2)^2}(-\sigma^2 + \frac{1}{n}\sum_{k=1}^n (x_k - \mu)^2). \end{pmatrix} \end{aligned}$$

Then

$$D(\log L_X(v))(v) = 0$$

iff

$$(\mu, \sigma^2) = (\overline{X_n}, \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2.$$

Now, we have to figure out if this is a min or a max.
Let $w = (\mu, \sigma^2) = (\overline{X_n}, \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2$. Then

$$\begin{aligned} D^2(\log L_X(w))(w) &= \begin{pmatrix} \partial_\mu \partial_\mu \log L_X(w) & \partial_{\sigma^2}\partial_{\sigma^2} \log L_X(w) \end{pmatrix} \\ &= \begin{pmatrix} \frac{-n}{\sigma^2} & \frac{-n}{2(\sigma^2)^2} \end{pmatrix}. \end{aligned}$$

Note that we evaluated the second derivative at the point w, and also the second derivative matrix is simpler since f is in the dual.
Let $u \in \mathbb{R}^2$. Then using the dot product

$$\langle D^2(\log L_X(w))u, u\rangle = \frac{-n}{\sigma^2}u_1 + \frac{-n}{2(\sigma^2)^2}u_2 \leq 0.$$

Hence, $D^2(\log L_X(w))$ is self-adjoint, and hence a negative operator (negative definite).
Therefore, w is a local max of the map L_X. Since v was arbitrary, w is a global max. Therefore, $w = (\overline{X_n}, \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2)$ is the MLE as required.

TODO finish Chapter 7.
8 / 13

Chapter 7 continued.
Bayesian
Risk or Average loss is a generalization of variance.
Median = $F^{-1}(1/2)$
L2 is a hilbert space
This

References

[1] Casella and Berger. *Statistical Inference.*