S721 HW9

John Koo

To save on typing, I will denote $\frac{\partial^k}{\partial x^k} f(x) = \partial_x^k f(x)$.

6.1

Note that $x^2 = |x|^2$. Therefore, $f(x \mid \sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-t^2/2\sigma^2}$. Then we can let $g(t \mid \sigma^2) = f(x \mid \sigma^2)$ and $h(x) = 1$.

6.2

$f(\overrightarrow{x} \mid \theta) = \prod_i^n e^{i\theta - x_i} 1_{(i\theta,\infty)}(x_i) = e^{\theta \sum_i i} e^{-\sum_i x_i} \prod_i 1_{(i\theta,\infty)}(x_i)$

Since one of the $x_i = \min_i x_i$, $e^{\theta \sum_i x_i} \prod_i 1_{(i\theta,\infty)}(x_i)$ is a function of $\min_i x_i/i$. The remaining part does not depend on θ, so we can write it as $h(\overrightarrow{x}) = e^{-\sum_i x_i}$.

6.3

$f(\overrightarrow{x} \mid \mu, \sigma) = = \prod_i^n \sigma^{-1} e^{-(x-\mu)/\sigma} 1_{(\mu,\infty)}(x_i)$
$= (\frac{e^{\mu/\sigma}}{\sigma})^n e^{-\sum_i x_i/\sigma} 1_{(\mu,\infty)}(x_{(1)})$ since we only need the smallest value to be greater than the lower bound μ.

Then letting $t_1 = \sum_i x_i$ and $t_2 = x_{(1)}$, we get $g(t_1, t_2 \mid \mu, \sigma) = (\frac{e^{\mu/\sigma}}{\sigma})^n e^{-t_1/\sigma} 1_{(\mu,\infty)}(t_2)$ and $h(\overrightarrow{x}) = 1$.

Therefore, $\overrightarrow{T} = \begin{bmatrix} \sum_i X_i \\ X_{(1)} \end{bmatrix}$

6.5

$f(\overrightarrow{x} \mid \theta) = \prod_i^n \frac{1}{2i\theta} 1_{(-i(\theta-1),i(\theta+1))}(x_i) = (2\theta)^{-n} \prod_i^n i^{-1} 1_{(-i(\theta-1),i(\theta+1))}(x_i)$

We can see that $\min_i x_i/i \geq -(\theta - 1)$ and $\max_i x_i/i \leq \theta + 1$, so this becomes:

$(2\theta)^{-n} I(\min_i x_i/i \geq -(\theta-1) I(\max_i x_i/i \leq \theta+1) \prod_i^n i^{-1}$

Then letting the first three terms be g and the last therm that doesn't depend on θ be h, we can see that

$T = \begin{bmatrix} \min_i x_i/i \\ \max_i x_i/i \end{bmatrix}$

6.6

$f(\overrightarrow{x} \mid \alpha, \beta) = \prod_i^n \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta} = (\Gamma(\alpha)\beta^\alpha)^{-n} (\prod_i^n x_i)^{\alpha-1} e^{-\sum_i x_i/\beta}$

Then letting $T = \begin{bmatrix} \prod_i x_i \\ \sum_i x_i \end{bmatrix}$, we get $(\Gamma(\alpha)\beta^\alpha)^{-n} t_1^{\alpha-1} e^{-t_2/\beta}$, which we can set to $g(T)$, letting $h(\overrightarrow{x}) = 1$.

7.6

Part a

$f(x \mid \theta) = theta^n \prod_i^n x_i^{-2} 1_{[\theta,\infty)}(x_i)$

Since we are only bounded to the left, we can pull out the indicator function using $x_{(1)}$, so we get:

$(\theta^n 1_{[\theta,\infty)}(x_{(1)}))(\prod_i^n x_i^{-2})$ and now we can see that the second part doesn't depend on θ so we can set it to $h(\overrightarrow{x})$. For the first part we can set $T = X_{(1)}$ to get $g(T \mid \theta)$.

7.10

Part a

We are given the distribution, so we need to derivate w.r.t. x to get the density.

$f(x \mid \alpha, \beta) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1}$ for $x \in [0, \beta]$ and 0 otherwise.

So the joint density is $f(\overrightarrow{x} \mid \alpha, \beta) = (\frac{\alpha}{\beta^\alpha})^n \prod_i^n x_i^{\alpha-1} 1_{[0,\beta]}(x_i)$

Since our support has two bounds, we need to care about both the max and min, so $\prod_i^n 1_{[0,\beta]}(x_i)$ becomes $1_{[0,\infty)}(x_{(1)})1_{(-\infty,\beta]}(x_{(n)})$. So the joint density is:
$(\frac{\alpha}{\beta^\alpha})^n 1_{[0,\infty)}(x_{(1)})1_{(-\infty,\beta]}(x_{(n)}) \prod_i^n x_i^{\alpha-1}$

The first, third, and fourth terms depend on α or β, so those go into our $g(T)$, so $h(\overrightarrow{x}) = 1_{[0,\infty)}(x_{(1)})$. Then we can just let $T = \begin{bmatrix} \prod_i^n X_i \\ X_{(n)} \end{bmatrix}$.

7.19

Part a

$f(\overrightarrow{y} \mid \overrightarrow{x}, \beta, \sigma^2) = \prod_i^n (2\pi\sigma^2)^{-1/2} e^{-(y_i-\beta x_i)^2/2\sigma^2} = (2\pi\sigma^2)^{-n/2} e^{-\sum_i y_i^2/2\sigma^2 + \beta \sum_i x_i y_i/\sigma^2 - \beta^2 \sum_i x_i^2/2\sigma^2}$

Here we have to set $h = 1$ since we cannot separate out β or σ^2.

We can see that the terms that depend on $\overrightarrow{y}$ are $\sum_i y_i^2$ and $\sum_i x_i y_i$. Therefore,

$T = \begin{bmatrix} \sum_i Y_i^2 \\ \sum_i x_i Y_i \end{bmatrix}$

11.6

Part a

We have $f_i(\overrightarrow{y}_i) = \prod_j^{n_i} (2\pi\sigma^2)^{-1/2} e^{(y_{ij}-\theta_i)^2/2\sigma^2}$

The full joint density is then $f = \prod_i^k f_i = \prod_i^k \prod_j^{n_i} (2\pi\sigma^2)^{-1/2} e^{(y_{ij}-\theta_i)^2/2\sigma^2}$
$= \prod_i^k (2\pi\sigma^2)^{-n_i/2} e^{-\sum_j (y_{ij}-\theta_i)^2/2\sigma^2}$

$= (2\pi\sigma^2)^{-\sum_i n_i/2} e^{-\sum_i \sum_j (y_{ij}-\theta_i)^2/2\sigma^2}$
$= (2\pi\sigma^2)^{-\sum_i n_i/2} \exp\Bigg(-\sum_i \sum_j y_{ij}^2 + \sum_i \theta_i \sum_j y_{ij}/\sigma^2 - \sum_i \sum_j \theta^2/2\sigma^2 \Bigg)$

So we have $t_i = \sum_j y_{ij}$ and $t_{k+1} = \sum_i \sum_j y_{ij}^2$.

But there exists a 1-1 mapping between $\sum_j Y_{ij}$ and $\bar{Y}_i$ as well as between $\sum_i \sum_j Y_{ij}^2$ and S^2.

11.35

Part a

We wish to minimize $\sum_i \epsilon_i^2 = \sum_i (y_i - \theta x_i^2)^2$, which can be done by derivating w.r.t. θ and then setting to 0. So we get:

$0 = \sum_i -2x_i^2(y_i - \theta x_i^2)$
$\implies 0 = \sum_i x_i^2 y_i - \theta \sum_i x_i^4$
$\implies \hat{\theta} = \frac{\sum_i x_i^2 Y_i}{\sum_i x_i^4}$

Part b

$Y_i \sim \mathcal{N}(\theta x_i^2, \sigma^2)$, so
$\mathcal{L}(\theta) = \prod_i^n (2\pi\sigma^2)^{-1/2} e^{(y_i - \theta x_i^2)^2/2\sigma^2}$

Then $\ell(\theta) = -\frac{n}{2} \log 2\pi\sigma^2 - \sum_i (y_i - \theta x_i^2)^2/2\sigma^2$

The first term doesn't depend on θ, and the second term is just the same as part (a) except with a constant, so derivating w.r.t. θ and setting to 0 will provide the same solution, $\hat{\theta} = \frac{\sum_i x_i^2 Y_i}{\sum_i x_i^4}$

Part c

$g(\theta) = \theta$, so $g'(\theta) = 1$ and $(g'(\theta))^2 = 1$.

$\partial_\theta \ell = \sum_i (y_i - \theta x_i^2) x_i^2/\sigma$,
so $\partial_\theta^2 \ell = -\sum_i x_i^4/\sigma^2$.
Since this doesn't depend on θ, the negative of its expected value is just $\sum_i x_i^4/\sigma^2$.

Then the CRLB is just $\frac{\sigma^2}{\sum_i x_i^4}$.

On the other hand, $Var(\frac{\sum_i x_i^2 Y_i}{\sum_i x_i^4}) = Var(\sum_i \frac{x_i^2 Y_i}{\sum_j x_j^4})$
$= \sum_i \frac{1}{(\sum_j x_j^4)^2} x_i^4 Var(Y_i)$
$= \frac{\sum_i x_i^4}{(\sum_j x_j^4)^2} n\sigma^2$
$\frac{n\sigma^2}{n \sum_i x_i^4} = \frac{\sigma^2}{\sum_i x_i^4}$ which is already the CRLB.

Which is just the CRLB.

11.37

Part a

Similar to 11.35b, $\ell = -\frac{n}{2}\log 2\pi\sigma^2 - \sum_i (y_i - \beta x_i)^2/2\sigma^2$

So $0 = \sum_i x_i(y_i - \beta x_i)/\sigma^2$
$\implies 0 = \sum_i x_i y_i - \beta \sum_i x_i^2$
$\implies \hat{\beta} = \frac{\sum_i x_i y_i}{\sum_i x_i^2}$

Part b

From part a, $\ell'(\beta) = \sum_i x_i(y_i - \beta x_i)/\sigma^2 = \sum_i x_i y_i/\sigma^2 - \beta \sum_i x_i^2/\sigma^2$

Then $\ell''(\beta) = -\sum_i x_i^2/\sigma^2$, which doesn't depend on β, so the negative of its expected value is just $\sum_i x_i^2/\sigma^2$.

Therefore, the CRLB is $\frac{\sigma^2}{\sum_i x_i^2}$

Part c

Similar to 11.35c, $Var(\frac{\sum_i x_i Y_i}{\sum_i x_i^2}) = \frac{\sigma^2}{\sum_i x_i^2}$, which is precisely the CRLB. So the MLE is the best estimator.

11.38

Part c

In HW6, we did parts a and b, so we know that the MLE is $\hat{\theta} = \frac{\sum_i Y_i}{\sum_i x_i}$.

We also computed its variance and obtained $Var(\hat{\theta}) = \frac{\theta}{\sum_i x_i}$.

In addition, we computed the log likelihood and its first derivative: $\ell'(\theta) = -\sum_i x_i + \frac{1}{\theta}\sum_i y_i$

Derivating again w.r.t. θ gives us $-\frac{\sum_i y_i}{\theta^2}$.

Then $-E[-\frac{\sum_i Y_i}{\theta^2}] = \frac{1}{\theta^2}\sum_i E[Y_i]$
$= \frac{1}{\theta^2}\sum_i \theta x_i$
$= \frac{\sum_i x_i}{\theta}$
which is just the CRLB.