HW5 solutions

1. (a) If $W(X)$ is an unbiased estimator of $\frac{1}{\theta}$, then

$E[W(X)] = \frac{1}{\theta}$ so the numerator of the CRLB is $\left(\frac{d}{d\theta}\left[\frac{1}{\theta}\right]\right)^2 = \left(-\frac{1}{\theta^2}\right)^2 = \frac{1}{\theta^4}$.

Since $X_1, \dots, X_n$ are iid random variables, the denominator of the CRLB is

$$n\,E\left[\left\{\frac{d}{d\theta}\ln f(x|\theta)\right\}^2\right] = n\,E\left[\left\{\frac{d}{d\theta}[\ln\theta - \theta X]\right\}^2\right]$$

$$= n\,E\left[\left\{\frac{1}{\theta} - X\right\}^2\right]$$

$$= n\,E\left[\left(X - \frac{1}{\theta}\right)^2\right]$$

$$= n\,\mathrm{Var}[X] \qquad \text{since } E[X] = \frac{1}{\theta} \text{ if } X \text{ is exponential with rate parameter } \theta$$

$$= n \cdot \frac{1}{\theta^2}.$$

So, the CRLB on the variance of an unbiased estimator of $\frac{1}{\theta}$ is

$$\frac{1/\theta^4}{n \cdot 1/\theta^2} = \boxed{\frac{1}{n\theta^2}}.$$

(b) The log likelihood function for estimating θ is

$$\ell(\theta) = \ln f(\underline{x}|\theta) = \ln \prod_{i=1}^{n} f(x_i|\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta) = \sum_{i=1}^{n}\{\ln\theta - \theta x_i\}$$

$$= n\ln\theta - \theta\sum_{i=1}^{n} x_i$$

So $\ell'(\theta) = \frac{n}{\theta} - \sum x_i$ and the MLE of θ is $\frac{1}{\bar{X}}$ since

$\ell'(\theta) = 0 \Rightarrow \frac{n}{\theta} = \sum x_i \Rightarrow \hat{\theta} = \frac{n}{\sum x_i} = \frac{1}{\bar{X}}$

sign of ℓ': $+$ | $-$

so $\ell'(\theta) > 0$ when $\theta < \hat{\theta}$ and $\ell'(\theta) < 0$ when $\theta > \hat{\theta}$.

By the invariance property of the MLE, the MLE of $\frac{1}{\theta}$ is $\frac{1}{\hat{\theta}} = \boxed{\bar{X}}$.

It is an unbiased estimator since $E[\bar{X}] = E[X_1] = \frac{1}{\theta}$ which satisfies the CRLB since $\mathrm{Var}[\bar{X}] = \frac{\mathrm{Var}[X_1]}{n} = \frac{1/\theta^2}{n} = \frac{1}{n\theta^2}$.

---

2. (a) $$\frac{f(\underline{x}|\mu,\sigma^2)}{f(y|\mu,\sigma^2)} = \frac{\cancel{(2\pi\sigma^2)^{-n/2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}}{\cancel{(2\pi\sigma^2)^{-n/2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2}} = \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2}}$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2 + \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2}$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i^2-2\mu x_i+\mu^2) + \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i^2-2\mu y_i+\mu^2)}$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2 + \frac{\mu}{\sigma^2}\sum x_i + \cancel{\frac{n\mu^2}{2\sigma^2}} + \frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 - \frac{\mu}{\sigma^2}\sum_{i=1}^n y_i - \cancel{\frac{n\mu^2}{2\sigma^2}}}$$

$$= e^{\frac{\mu}{\sigma^2}\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right)}$$

(~~b~~d) The joint pdf of $X_1,\dots,X_n$ is

$$f(\underline{x}|\mu,\sigma^2) = (2\pi)^{-n/2}(\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)}$$

$$= h(x_1,\dots,x_n)\, g\left(\sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i \,\Big|\, \mu,\sigma^2\right)$$

where $h(x_1,\dots,x_n) = (2\pi)^{-n/2}$ and $g\left(\sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i \,|\, \mu,\sigma^2\right) = (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\left(\sum x_i^2 - 2\mu\sum x_i + n\mu^2\right)}$

so, by the Factorization Theorem, $\left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$ is sufficient for ~~(μ,σ^2)~~ μ

in parts (b) and (d). Note that in both cases σ^2 is a function of μ.

(b) If $\mu=\sigma^2$, then $f(\underline{x}|\mu) = (2\pi)^{-n/2}\mu^{-n/2} e^{-\frac{1}{2\mu}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)}$

$$= \left[(2\pi)^{-n/2} e^{\sum_{i=1}^n x_i}\right]\left[\mu^{-n/2} e^{-\frac{1}{2\mu}\sum_{i=1}^n x_i^2} e^{-\frac{n\mu}{2}}\right]$$

$$= h_2(x_1,\dots,x_n)\, g_2\left(\sum_{i=1}^n x_i^2 \,\Big|\, \mu\right)$$

where $h_2(x_1,\dots,x_n) = (2\pi)^{-n/2} e^{\sum x_i}$ and $g_2(\sum x_i^2|\mu) = \mu^{-n/2} e^{-\frac{1}{2\mu}\sum x_i^2} e^{-n\mu/2}$

so $\sum_{i=1}^n X_i^2$ is sufficient for μ by the Factorization Theorem.

Hence $\left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$ is not minimal sufficient since it is not a function of $\sum_{i=1}^n X_i^2$, as required by Definition L 10.2.

(c) If $\mu=\sigma^2$, then $\frac{f(\underline{x}|\mu)}{f(y|\mu)} = e^{\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right) - \frac{1}{2\mu}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right)} = C(\underline{x},\underline{y})$ ← constant w.r.t. μ

if and only if $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$. So, by Theorem L 10.4,

$\sum_{i=1}^n X_i^2$ is minimal sufficient for μ.

To see ~~this~~ the "if and only if" statement in more detail, suppose $\sum x_i^2 = \sum y_i^2$. Then $\frac{f(\underline{x}|\mu)}{f(\underline{y}|\mu)} = e^{\sum x_i - \sum y_i}$ does not depend on μ.

Conversely, suppose $\dfrac{f(\underline{x}|\mu)}{f(y|\mu)} = c(\underline{x},y)$ for all μ.

Then for $\mu = \frac{1}{2}$ and $\mu = \frac{1}{4}$, $c(\underline{x},y) = e^{\Sigma x_i - \Sigma y_i - (\Sigma x_i^2 - \Sigma y_i^2)} = e^{\Sigma x_i - \Sigma y_i - 2(\Sigma x_i^2 - \Sigma y_i^2)}$

$$\Rightarrow \Sigma x_i - \Sigma y_i - (\Sigma x_i^2 - \Sigma y_i^2) = \Sigma x_i - \Sigma y_i - 2(\Sigma x_i^2 - \Sigma y_i^2)$$

$$\Rightarrow \Sigma x_i^2 - \Sigma y_i^2 = 2(\Sigma x_i^2 - \Sigma y_i^2)$$

$$\Rightarrow \Sigma x_i^2 - \Sigma y_i^2 = 0 \Rightarrow \Sigma x_i^2 = \Sigma y_i^2 .$$

(d) If $\mu = \sigma$, then $\dfrac{f(\underline{x}|\mu)}{f(y|\mu)} = e^{\frac{1}{\mu}(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i) - \frac{1}{2\mu^2}(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2)} = c(\underline{x},y)$

if and only if $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$ and $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. So, by Theorem L10.4,

$\left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$ is minimal sufficient for μ.

To see this in more detail, suppose $\Sigma x_i^2 = \Sigma y_i^2$ and $\Sigma x_i = \Sigma y_i$. Then

$\dfrac{f(\underline{x}|\mu)}{f(y|\mu)} = e^0 = 1$ does not depend on μ.

Conversely, if $\dfrac{f(\underline{x}|\mu)}{f(y|\mu)} = c(\underline{x},y)$ for all μ, then

$$c(\underline{x},y) = \underset{\mu = 1/2}{\underset{\uparrow}{e^{2(\Sigma x_i - \Sigma y_i) - 2(\Sigma x_i^2 - \Sigma y_i^2)}}} = \underset{\mu = 1}{\underset{\uparrow}{e^{(\Sigma x_i - \Sigma y_i) - \frac{1}{2}(\Sigma x_i^2 - \Sigma y_i^2)}}} = \underset{\mu = 1/\sqrt{2}}{\underset{\uparrow}{e^{\sqrt{2}(\Sigma x_i - \Sigma y_i) - (\Sigma x_i^2 - \Sigma y_i^2)}}} .$$

Letting $A = \Sigma x_i - \Sigma y_i$ and $B = \Sigma x_i^2 - \Sigma y_i^2$, we get the system of equations

$$\underbrace{2A - 2B = \overbrace{A - \tfrac{1}{2}B}^{0 = (\sqrt{2}-1)A - \frac{1}{2}B}}_{A - \frac{3}{2}B = 0} \overbrace{= \sqrt{2}A - B}$$

so $A = \frac{3}{2}B \Rightarrow (\sqrt{2}-1)\frac{3}{2}B - \frac{1}{2}B = 0$

$\Rightarrow \left(\frac{3}{2}(\sqrt{2}-1) - \frac{1}{2}\right)B = 0$

$\Rightarrow B = 0$

$\Rightarrow A = \frac{3}{2} \cdot 0 = 0 .$

Thus, $\Sigma x_i - \Sigma y_i = 0 \Rightarrow \Sigma x_i = \Sigma y_i$

and $\Sigma x_i^2 - \Sigma y_i^2 = 0 \Rightarrow \Sigma x_i^2 = \Sigma y_i^2 .$

3. (a) $E[G_1] = 0 \cdot P(G_1=0) + 1 \cdot P(G_1=1) = P(G_1=1) = P(X_1>1)$

$$= 1 - P(X_1 \le 1)$$

$$= 1 - P(X_1=0) - P(X_1=1)$$

$$= 1 - p - p(1-p)$$

$$= 1(1-p) - p(1-p)$$

$$= (1-p)(1-p) = (1-p)^2.$$

(b) The joint pmf of $X_1, \dots, X_n$ is

$$f(\underline{x}\mid p) = \prod_{i=1}^{n} f(x_i\mid p) = \prod_{i=1}^{n} p(1-p)^{x_i} I_{\mathbb{Z}^*}(x_i) = p^n (1-p)^{\sum_{i=1}^n x_i} \prod_{i=1}^n I_{\mathbb{Z}^*}(x_i)$$

with $\mathbb{Z}^*$ being the set of all nonnegative integers

$$= h(x_1,\dots,x_n)\, g\left(\sum_{i=1}^n x_i \mid p\right)$$

where $h(x_1,\dots,x_n) = \prod_{i=1}^n I_{\mathbb{Z}^*}(x_i)$

and $g\left(\sum_{i=1}^n x_i \mid p\right) = p^n (1-p)^{\sum_{i=1}^n x_i}$

So $\sum_{i=1}^n X_i$ is sufficient by the Factorization Theorem.

(c) Since $X_1,\dots,X_n$ are independent geometric random variables ~~with~~ representing the number of failures before a success, $\sum_{i=1}^n X_i$ is a negative binomial random variable representing the number of failures before n successes. The pmf of $\sum_{i=m}^n X_i$ is

$$P\left(\sum_{i=m}^n X_i = y\right) = \underbrace{\binom{n-m+y}{y}}_{\substack{\text{\# of ways of arranging} \\ \text{(y F's) among (n-m S's and y F's)}}} p^{n-m+1}(1-p)^y \text{ for } m \le n.$$

So

$$P\left(X_1 = x \mid \sum_{i=1}^n X_i = t\right) = \frac{P(X_1 = x \text{ and } X_1+X_2+\dots+X_n = t)}{P\left(\sum_{i=1}^n X_i = t\right)}$$

$$= \frac{P(X_1 = x \text{ and } X_2+\dots+X_n = t-x)}{P\left(\sum_{i=1}^n X_i = t\right)}$$

$$= \frac{P(X_1 = x)\, P(X_2 + \dots + X_n = t - x)}{P(\sum_{i=1}^{n} X_i = t)}$$ by independence of X_1 and $(X_2, \dots, X_n)$

$$= \frac{p(1-p)^x \binom{n-2+t-x}{t-x} p^{n-1}(1-p)^{t-x}}{\binom{n-1+t}{t} p^n (1-p)^t}$$

$$= \frac{\binom{n-2+t-x}{t-x} \cancel{p^n} \cancel{(1-p)^t}}{\binom{n-1+t}{t} \cancel{p^n} \cancel{(1-p)^t}} = \boxed{\frac{\binom{n-2+t-x}{t-x}}{\binom{n-1+t}{t}}} \text{ for } x = 0, \dots, t.$$

$$= \begin{cases} \dfrac{(n-1)[t \cdots (t-x+1)]}{(n+t-1)\cdots(n+t-x-1)} & \text{if } t \geq 1 \\ 1 & \text{for } x=0 \text{ if } t=0 \end{cases}.$$

(d) By the Rao-Blackwell Theorem, $\phi(\Sigma X_i)$ is unbiased for $(1-p)^2$ and uniformly better than G_1 where $\phi(t) = E[G_1 \mid \Sigma X_i = t]$.

So $\phi(t) = E[G_1 \mid \Sigma X_i = t] = P(G_1 = 1 \mid \Sigma X_i = t)$

$$= P(X_1 > 1 \mid \Sigma X_i = t)$$

$$= 1 - P(X_1 = 0 \mid \Sigma X_i = t) - P(X_1 = 1 \mid \Sigma X_i = t)$$

$$= 1 - \frac{\binom{n-2+t}{t}}{\binom{n-1+t}{t}} - \frac{\binom{n-3+t}{t-1}}{\binom{n-1+t}{t}}.$$

Then $\phi(\Sigma X_i)$
$$= \boxed{\begin{cases} 1 - \dfrac{\binom{n-2+\Sigma X_i}{\Sigma X_i}}{\binom{n-1+\Sigma X_i}{\Sigma X_i}} - \dfrac{\binom{n-3+\Sigma X_i}{\Sigma X_i - 1}}{\binom{n-1+\Sigma X_i}{\Sigma X_i}} & \text{if } \Sigma X_i > 1 \\ 0 & \text{if } \Sigma X_i \leq 1 \end{cases}}$$

$$= \begin{cases} 1 - \dfrac{n-1}{n+\Sigma X_i - 1} - \dfrac{(n-1)\Sigma X_i}{(n+\Sigma X_i - 2)(n+\Sigma X_i - 1)} & \text{if } \Sigma X_i > 1 \\ 0 & \text{if } \Sigma X_i \leq 1 \end{cases}.$$