Problem 1. Suppose that $X_1, \ldots, X_n$ are drawn i.i.d from a distribution with density

$$p_\theta(x) = \theta e^{-x\theta}$$

where $x > 0$ and $\theta > 0$.

(a) Find the MLE

(b) Find the Fisher Information

(c) Construct an asymptotic $1 - \alpha$ confidence interval for $\theta$.

(d) Construct an asymptotic $1 - \alpha$ confidence interval for $\log(\theta)$

(e) Find the pdf of $Y = 1/X$

Solution 1. We derive each of the results below:

(a) The likelihood of $\theta$ is calculated as follows.

$$L(\theta) = \mathbb{P}(X_1, \ldots, X_n|\theta)$$

$$= \prod_{i=1}^{n} \mathbb{P}(X_i|\theta)$$

$$= \prod_{i=1}^{n} \theta e^{-X_i \theta}$$

$$= \theta^n e^{-\theta \sum_{i=1}^{n} X_i}.$$

Thus, the log likelihood is,

$$\ell(\theta) = \log(L(\theta)) = n \log(\theta) - \theta \sum_{i=1}^{n} X_i.$$

Differentiate with respect to $\theta$ and set the derivative to 0 to get the MLE,

$$\frac{d\ell(\theta)}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} X_i.$$

Hence,

$$\frac{d\ell(\theta)}{d\theta} = 0 \Rightarrow \frac{n}{\theta} - \sum_{i=1}^{n} X_i = 0 \Rightarrow \hat{\theta}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}.$$
This is indeed a maximum since the second derivative is always negative,
\[
\frac{d^2\ell(\theta)}{d\theta^2} = -\frac{n}{\theta^2} < 0.
\]
Hence, the MLE estimator for \(\theta\) is \(\hat{\theta}_{\text{MLE}} = 1/\bar{X}\).

(b)
\[
\mathcal{I}_n(\theta) = -\mathbb{E}\left(\frac{d^2\ell(\theta)}{d\theta^2}\right)
= -\mathbb{E}\left(-\frac{n}{\theta^2}\right)
= \frac{n}{\theta^2}.
\]

By the asymptotic normality of MLE estimators, we have \(\hat{\theta}_{\text{MLE}} \overset{d}{\rightarrow} \mathcal{N}\left(\theta, \frac{1}{n}\right)\). Hence, the \(1 - \alpha\) asymptotic confidence interval for \(\theta\) is
\[
\theta \in \left[\hat{\theta}_{\text{MLE}} - z_{\alpha/2} \frac{\theta}{\sqrt{n}}, \hat{\theta}_{\text{MLE}} + z_{\alpha/2} \frac{\theta}{\sqrt{n}}\right].
\]

(d) Let \(p := \log \theta = g(\theta)\). Since \(g\) is continuous, by the equivalence principle of MLE, \(\hat{p}_{\text{MLE}} = \log(\hat{\theta}_{\text{MLE}})\). Then \(g'(\theta) = 1/\theta\). By the Delta method, since \(\hat{\theta}_{\text{MLE}} \overset{d}{\rightarrow} \mathcal{N}(\theta, \frac{\theta^2}{n})\), we have
\[
\hat{p}_{\text{MLE}} = \log(\hat{\theta}_{\text{MLE}}) \overset{d}{\rightarrow} \mathcal{N}(\log(\theta), \frac{\theta^2}{n} (g'(\theta)^2) = \mathcal{N}(\log(\theta), \frac{1}{n}).
\]
Hence, the \(1 - \alpha\) asymptotic confidence interval for \(\log(\theta)\) is
\[
\log(\theta) \in \left[\log(\hat{\theta}_{\text{MLE}}) - z_{\alpha/2} \frac{1}{\sqrt{n}}, \log(\hat{\theta}_{\text{MLE}}) + z_{\alpha/2} \frac{1}{\sqrt{n}}\right].
\]

(e)
\[
F_Y(y) = \mathbb{P}(Y \leq y)1_{y>0}
= \mathbb{P}(1/X \leq y)1_{y>0}
= \mathbb{P}(X \geq 1/y)1_{y>0}
= [1 - \mathbb{P}(X \leq 1/y)]1_{y>0}
= [1 - F_X(1/y)]1_{y>0}
\]

Thus, the PDF of \(Y\) is
\[
f_Y(y) = \frac{dF_Y(y)}{dy}
= -f_X(1/y)(-\frac{1}{y^2})1_{y>0}
= \frac{\theta}{y^2}e^{\theta/y}1_{y>0}.
\]
Problem 2. Let $X_1, X_2, \ldots, X_n$ be I.I.D random variables such that $\mathbb{E}(X_i) = \mu < +\infty$ and $\text{Var} X_i = \sigma^2 < +\infty, \forall i \in [n]$. State and prove the weak law of large numbers (WLLN) for $X_1, \ldots, X_n$.

Solution 2. We derive each of the results below:

(a) Let $X_1, X_2, \ldots, X_n$ be random variables such that $\mathbb{E}(X_i) = \mu < +\infty$ and $\text{Var} X_i = \sigma^2 < +\infty, \forall i \in [n]$. The Weak Law of Large Numbers states that:

$$\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{p} \mu$$

Proof. We note that

$$\mathbb{E} \left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n} (n\mu) = \mu$$

So we have:

$$\mathbb{P} \left( \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| > \epsilon \right) \leq \frac{\text{Var} \left( \frac{1}{n} \sum_{i=1}^{n} X_i \right)}{\epsilon^2} \quad \text{(Using Chebychev's inequality)}$$

$$= \frac{\frac{1}{n^2} (n\sigma^2)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \rightarrow 0 \text{ as } n \rightarrow \infty$$

\qed
Problem 3. Let $X_1, \ldots, X_n \sim N(\theta, 1)$. Let $\theta$ have $N(0, 1)$ prior. That is,

$$p(\theta) = \frac{1}{\sqrt{\pi}} e^{-\frac{\theta^2}{2}}$$

(a) Find the Bayes estimator (using squared error loss).

(b) Find the risk of the estimator.

Solution 3. Define the following:

(a) We claim that $E(X) = \mu$. To derive the Bayes estimator of $p$ we need to calculate the posterior mean (under squared loss). This is done as follows:

Proof.

$$P(\theta \mid x_1, \ldots, x_n) \propto L(x_1, \ldots, x_n \mid \theta) \pi(\theta)$$

$$= \left[ \prod_{i=1}^{n} f_{X_i}(x_i) \right] \frac{1}{\sqrt{\pi}} \exp \left( -\frac{\theta^2}{2} \right)$$

$$= \left[ \prod_{i=1}^{n} \frac{1}{\sqrt{\pi}} \exp \left( -\frac{(x_i - \theta)^2}{2} \right) \right] \frac{1}{\sqrt{\pi}} \exp \left( -\frac{\theta^2}{2} \right)$$

$$\propto \left[ \exp \sum_{i=1}^{n} \left( -\frac{(x_i - \theta)^2}{2} \right) \right] \exp \left( -\frac{\theta^2}{2} \right)$$

$$\propto \exp \left[ -\frac{1}{2} \sum_{i=1}^{n} \left( -2\theta x_i + \theta^2 \right) - \frac{1}{2} \theta^2 \right]$$

$$= \exp \left[ -\frac{1}{2} \left( -2\theta (n \bar{x}_n) + n\theta^2 \right) - \frac{1}{2} \theta^2 \right]$$

$$= \exp \left[ -\frac{\theta^2}{2} (n + 1) + \theta (n \bar{x}_n) \right]$$

Now let $(\theta')^2 = (n + 1)^{-1}$ and let $\theta' = (\theta')^2 (n \bar{x}_n)$. We can then continue:

$$P(\theta \mid x_1, \ldots, x_n) \propto \exp \left[ -\frac{\theta^2}{2} (n + 1) + \theta (n \bar{x}_n) \right]$$

$$= \exp \left[ -\frac{\theta^2}{2 (\theta')^2} + \frac{\theta \theta'}{(\theta')^2} \right]$$

$$\propto \exp \left[ -\frac{\theta^2}{2 (\theta')^2} + \frac{\theta \theta'}{(\theta')^2} - \frac{(\theta')^2}{2 (\theta')^2} \right]$$

$$\propto \exp \left[ -\frac{(\theta - \theta')^2}{2 (\theta')^2} \right]$$

$$\sim \mathcal{N}(\theta', (\theta')^2)$$

$$= \mathcal{N}(\frac{n \bar{x}_n}{n + 1}, \frac{1}{n + 1})$$
(b) The Bayes estimator (using squared error loss) is the posterior mean i.e.

\[ \hat{\theta}_{\text{BAYES}} = \mathbb{E}(\theta | X_1, \ldots, X_n) \]
\[ = \frac{n \bar{X}_n}{n + 1} \]

(c) The risk of the Bayes estimator \( \hat{\theta}_{\text{BAYES}} \) is the mean squared error (MSE) (under square loss)

\[ \text{MSE}(\hat{\theta}_{\text{BAYES}}) = (\text{Bias}(\hat{\theta}_{\text{BAYES}}))^2 + \text{Var}(\hat{\theta}_{\text{BAYES}}) \]

Now we calculate the Bias and variance separately

\[ \mathbb{E}(\hat{\theta}_{\text{BAYES}}) = \mathbb{E} \left( \frac{n \bar{X}_n}{n + 1} \right) \]
\[ = \frac{n}{n + 1} \mathbb{E}(\bar{X}_n) \]
\[ = \frac{n}{n + 1} \mathbb{E}(X_1) \]
\[ = \frac{n \theta}{n + 1} \]

So we have that \( \text{Bias}(\hat{\theta}_{\text{BAYES}}) = \mathbb{E}(\hat{\theta}_{\text{BAYES}}) - \theta = -\frac{\theta}{n + 1} \). And now the variance

\[ \text{Var}(\hat{\theta}_{\text{BAYES}}) = \left( \frac{n}{n + 1} \right)^2 \frac{1}{n} \text{Var}(X_1) \]
\[ = \left( \frac{n}{n + 1} \right)^2 \frac{1}{n} \]
\[ = \frac{n}{(n + 1)^2} \]

So finally the MSE is:

\[ \text{MSE}(\hat{\theta}_{\text{BAYES}}) = (\text{Bias}(\hat{\theta}_{\text{BAYES}}))^2 + \text{Var}(\hat{\theta}_{\text{BAYES}}) \]
\[ = \frac{\theta^2 + n}{(n + 1)^2} \]
Problem 4. Let $X_1, \ldots, X_n \sim F$ i.i.d for some (continuously and strictly increasing) cdf $F$. Let $\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$.

(a) Use Hoeffding’s inequality to show that, for any $t > 0$,

$P(|\hat{F}_n(x) - F(x)| > t) \to 0.$

(b) Let $X \sim F$ be another observation from $X$. Let $T = \hat{F}_n(X)$. Find the mean and variance of $T$. Hint: Consider conditioning on $X_1, \ldots, X_n$.

Solution 4. (a) Let $Y_i = I(X_i \leq x)$. Then $Y_i$ are also i.i.d with $0 \leq Y_i \leq 1$. Notice that $E(Y) = E(\hat{F}_n(x)) = F(x)$. By Hoeffding’s inequality,

$P(|\hat{F}_n(x) - F(x)| > t) = P(|\bar{Y} - F(x)| > t)
\leq 2 \exp\left(\frac{-2nt^2}{\sum_{i=1}^{n}(1-0)^2}\right)
= 2 \exp(-2nt^2) \to 0 \quad \text{as } n \to \infty$

Therefore

$P(|\hat{F}_n(x) - F(x)| > t) \to 0$

(b)

$E(T) = E(E(T|X_1, \ldots, X_n))$

$= \frac{1}{n} \sum_{i=1}^{n} E(E(I(X_i \leq X)|X_i))$

$= E(E(I(X_i \leq X)|X_i))$

$= E(1 - F(X_i)) \quad \text{Since } F \text{ is continuous}
= \frac{1}{2} \quad \text{Since } F(X_i) \sim U[0, 1] \text{ (See claim for the proof)}$

$Var(T) = E(Var(T|X_1, \ldots, X_n)) + Var(E(T|X_1, \ldots, X_n))$

And

$E\left(Var(T|X_1, \ldots, X_n)\right) = \frac{1}{n} E\left(Var(I(X_i \leq X)|X_i)\right)$

$= \frac{1}{n} E\left(E(I^2(X_i \leq X)|X_i) - E^2(I(X_i \leq X)|X_i)\right)$

$= \frac{1}{n} E\left(E(I(X_i \leq X)|X_i) - E^2(I(X_i \leq X)|X_i)\right)$

$= \frac{1}{n} E\left((1 - F(X_i)) - (1 - F(X_i))^2\right)$

$= \frac{1}{n} E\left(- F^2(X_i) + F(X_i)\right)$

$= \frac{1}{n} \left(-\frac{1}{3} + \frac{1}{2}\right)$

$= \frac{1}{6n}$
\[
\text{Var}(\mathbb{E}(T|X_1, \ldots, X_n)) = \text{Var}(\mathbb{E}(I_{X_i \leq X_i}|X_i))
\]
\[
= \text{Var}(1 - F(X_i)) = \frac{1}{12}
\]

Thus
\[
\text{Var}(T) = \frac{1}{6n} + \frac{1}{12}
\]
Appendix: Supporting Lemmas

Here we prove some lemmas from the lecture notes that we use repeatedly in the solutions.

**Lemma 0.1** (CDF transform of RV is Uniform). If $X \sim F$ for a continuously and strictly increasing $F$, then $F(X) \sim \text{Unif}(0,1)$.

**Proof.** Let $Z = F(X)$, then for any $z \in [0,1]$,

\[
F_Z(z) = P(Z \leq z) = P(F(X) \leq z) = P(X \leq F^{-1}(z)) = F(F^{-1}(z)) = z
\]

Hence $Z \sim \text{Unif}(0,1)$, as required. \qed