Solutions to Practice Final Exam

(1) The cdf is

\[ F_X(x) = \begin{cases} 
0 & x < 0 \\
\frac{x}{2} & 0 \leq x < 1 \\
\frac{3}{2} & 1 \leq x < 3 \\
\frac{x^2}{2} + \frac{3-x}{2} & 3 \leq x < 4 \\
1 & x \geq 4. 
\end{cases} \]

(2) We know that

\[ \rho(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}} = \frac{\text{Cov}(X_i, X_j)}{\sigma^2} \]

so that \( \text{Cov}(X_i, X_j) = \sigma^2 \rho \). Hence,

\[
\text{Var}(\sum_i X_i) = \sum_i \text{Var}(X_i) + 2 \sum_{i<j} \text{Cov}(X_i, X_j)
\]

\[
= \sum_i \text{Var}(X_i) + 2 \sum_{i<j} \text{Cov}(X_i, X_j)
\]

\[
= \sum_i \text{Var}(X_i) + 2 \sum_{i<j} \sigma^2 \rho
\]

\[
= n\sigma^2 + n(n-1)\sigma^2 \rho.
\]

Since, \( \text{Var}(\sum_i X_i) \geq 0 \), conclude that

\[ n\sigma^2 + n(n-1)\sigma^2 \rho \geq 0 \]

and hence \( \rho \geq -1/(n-1) \).

(3) Let \( Z \sim N(0, 1) \). Then,

\[
\psi = E(Y_i) = (1 \times P(X > 0)) + (-1 \times P(X < 0))
\]

\[
= P(X > 0) - P(X < 0)
\]

\[
= (1 - P(X < 0)) - P(X < 0)
\]

\[
= 1 - 2P(X < 0)
\]

\[
= 1 - 2P(X - \theta < -\theta)
\]

\[
= 1 - 2\Phi(-\theta) \equiv g(\theta).
\]
(3a) The mle is $\hat{\psi} = g(\hat{\theta}) = g(\overline{X}) = 1 - 2\Phi(-\overline{X})$.

(3b) The estimated standard error of $\hat{\theta}$ is $se(\hat{\theta}) = 1/\sqrt{n}$. Now use the delta method: $gt(\theta) = 2\phi(-\theta) = 2\phi(\theta)$ and the estimated standard error of $\hat{\psi}$ is $se(\hat{\psi}) = se(\theta)|g'(\hat{\theta})| = 2\phi(\overline{X})/\sqrt{n}$. The approximate confidence interval is

$$1 - 2\Phi(-\overline{X}) \pm 2 \left(2\phi(\overline{X})/\sqrt{n}\right).$$

(3c) The variance of $\hat{\psi}$ is

$$Var(\hat{\psi}) = \frac{Var(Y_1)}{n}.$$

Now, $Var(Y_1) = E(Y_1^2) - \psi^2 = 1 - \psi^2$. The ARE is

$$\frac{se(\hat{\psi})}{se(\hat{\psi})} = \frac{2\phi(\theta)}{\sqrt{1 - \psi^2}} = \frac{\phi(\theta)}{\sqrt{\Phi(\theta)(1 - \Phi(\theta))}}.$$

(4) Note that $f(X_i; \theta)$ is 0 if $X_i > \theta$ or $X_i < -\theta$. The likelihood will be 0 if this happens for any $X_i$. So $\mathcal{L}(\theta)$ is non-zero only if $-\theta \leq X_i \leq \theta$ for all $X_i$. This is the same as $\max_i |X_i| \leq \theta$. On the other hand, $f(X_i; \theta) = 1/(2\theta)$ if $-\theta \leq X_i \leq \theta$. So,

$$\mathcal{L}_n(\theta) = \begin{cases} \left(\frac{1}{2\theta}\right)^n & \theta \geq \max_i |X_i| \\ 0 & \text{otherwise.} \end{cases}$$

(4b) $\hat{\theta} = \max_i |X_i|$.

(4c) First, $0 \leq \hat{\theta} \leq \theta$. For $c \in [0, \theta]$,

$$H(c) = P(\hat{\theta} \leq c)$$

2
\begin{align*}
&= P(\max_i |X_i| \leq c) \\
&= \prod_i P(|X_i| \leq c) \\
&= \prod_i P(-c \leq X_i \leq c) \\
&= \prod_i \int_{-c}^c f(x; \theta)dx \\
&= \prod_i \frac{c}{\theta} \\
&= \left(\frac{c}{\theta}\right)^n.
\end{align*}

The density is \( h(c) = n\theta^{-n}c^{n-1} \) for \( c \in [0, \theta] \).

(5a) \( \mathcal{L}_n(p) = p^Y (1 - p)^{n - Y} \) where \( Y = \sum_i X_i \) and \( \ell(p) = Y \log p + (n - Y) \log(1 - p) \). Setting \( \ell'(p) = 0 \) yields \( \hat{p} = Y/n \).

(5b) Let \( f(x; p) = p^x (1 - p)^{1-x} \). The Fisher information is
\[
I(p) = E\left(-\frac{\partial^2 \log f}{\partial p^2}\right) = \frac{1}{p(1 - p)}.
\]

(5c) \( \hat{p} \approx N(p, 1/(nI(p))) = N(p, p(1 - p)/n) \).

(5d) \( \text{MSE} = \text{bias}^2 + \text{var} = 0 + p(1 - p)/n = p(1 - p)/n \).

(6a) Let \( Q = \sum_i (Y_i - \beta x_i)^2 \). Setting \( dQ/d\beta = 0 \) gives \( \hat{\beta} = \frac{\sum_i x_i Y_i}{\sum_i x_i^2} \).

(6b) \( E(\hat{\beta}) = \frac{\sum_i x_i E(Y_i)/\sum_i x_i^2}{\beta} = \frac{\sum_i x_i x_i}{\sum_i x_i^2} = \beta \frac{\sum_i x_i^2/\sum_i x_i^2}{\beta} \) as long as \( \sum_i x_i^2 \neq 0 \) i.e. as long as \( x_i \neq 0 \) for some \( x_i \).
(6c) The variance is \( \text{Var}(\hat{\beta}) = \sum_i x_i^2 \text{Var}(Y_i)/(\sum_i x_i^2)^2 = \sigma^2 \sum_i x_i^2/(\sum_i x_i^2)^2 = \sigma^2/(\sum_i x_i^2) \). If \( \sum_i x_i^2 \to \infty \) then \( \text{MSE} \to 0 \) and then \( \hat{\beta} \to \beta \).

(7a) First, 

\[
F_Y(0) = P(Y \leq 0) = P(Y = 0) = P(X = 1) = \int_0^1 e^{-x^2} dx = 1 - e^{-1}.
\]

For \( 0 < y < 1 \), \( F_Y(y) = F_Y(0) \). For \( y \geq 1 \), \( F_Y(y) = P(Y \leq y) = P(Y = 0) + P(1 \leq Y \leq y) = P(Y = 0) + P(1 \leq X \leq y) = 1 - e^{-1} + \int_y^1 e^{-x^2} dx = 1 - e^{-1} + (1 - e^{-y}) = 2 - e^{-1} - e^{-y} \).

(7b) Write \( Y = r(X) \) where \( r(x) = 0 \) for \( 0 \leq x \leq 1 \) and \( r(x) = x \) for \( x \geq 1 \). So, 

\[
E(Y) = \int_0^\infty r(x)e^{-x} dx = \int_1^\infty xe^{-x} dx = \frac{2}{e}.
\]

(7c) Given \( X = x \), \( Y = r(X) \) is a point mass at \( r(x) \). So \( E(Y|X = x) = r(x) \). Hence, \( E(Y|X) = r(X) \).

(8a) Let \( X_{(1)} \) and \( X_{(n)} \) be the smallest and largest values. The likelihood is 0 unless \( \theta < X_{(1)} \) and \( 2\theta > X_{(n)} \) i.e. \( x_{(n)}/2 \leq \theta \leq X_{(1)} \). Thus,

\[
\mathcal{L}_n(\theta) = \begin{cases} 
\left(\frac{1}{\theta}\right)^n & \frac{x_{(n)}}{2} \leq \theta \leq X_{(1)} \\
0 & \text{otherwise}.
\end{cases}
\]

This is a decreasing function so \( \hat{\theta} = X_{(1)} \).

(8b) First, \( \theta \leq \hat{\theta} \leq 2\theta \). For \( c \in [\theta, 2\theta] \),

\[
H(c) = P(\hat{\theta} \leq c) = P(\min_i X_i \leq c) = 1 - P(\min_i X_i \geq c) = 1 - \prod_i P(X_i \geq c)
\]
\[
\begin{align*}
&= 1 - \prod_i \int_c^{2\theta} f(x; \theta) dx \\
&= 1 - \prod_i \frac{2\theta - c}{\theta} \\
&= 1 - \left(\frac{2\theta - c}{\theta}\right)^n.
\end{align*}
\]

The density is \( h(c) = (n/\theta)(2 - (c/\theta))^{n-1} \) for \( c \in [\theta, 2\theta] \).

(8c) For \( n = 2 \) the density is \( h(c) = (2/\theta)(2 - (c/\theta)). \) The mean is
\[
\int_{\theta}^{2\theta} c h(c) dc = \frac{4\theta}{3}.
\]

(9) Parts (i)-(iii) should be straightforward. For part (iv), note that \( Y = \sum X_i \sim \text{Poisson}(n\lambda) \). Let \( h(c) = P(\hat{\lambda} = c) \). Then, if \( nc \) is a nonnegative integer,
\[
h(c) = P(Y = nc) = \frac{e^{-n\lambda}(n\lambda)^{nc}}{(nc)!}.
\]
For part (v), note that \( se(\hat{\lambda}) = 1/\sqrt{nI(\hat{\lambda})} = \sqrt{\lambda}/n = \sqrt{X}/n. \) Also, \( \psi = e^{-\lambda} \equiv g(\lambda) \). So \( se(\hat{\psi}) = se(\hat{\lambda})|g'| = \sqrt{X}/ne^{-\lambda}. \) The confidence interval is
\[
e^{-\lambda} \pm 2 \sqrt{X}/ne^{-\lambda}.
\]

(10) Let \( Y = \sum X_i \). The posterior density is
\[
f(\lambda|X^n) \propto \frac{1}{\sqrt{\lambda}} e^{-n\lambda} \lambda^Y = e^{-n\lambda} \lambda^{Y-(1/2)}.
\]
This is Gamma \((Y + (1/2), 1/n)\). The mean of a Gamma \((\alpha, \beta)\) is (in this version) \( \alpha/\beta \). So the posterior mean is \( Y/n \).