ESE 524 - Homework 1

Solution Outline
Assigned date: 01/22/19
Due Date: 02/05/19
Total Points: 100

These solutions are meant to be sketches. For full solutions we encourage you to fill in the details on your own or ask the TA in the office hours.

(1) **Transformation of Random Variables**

a) Let $X \sim \text{Unif}[0, 1]$ be a uniform random variable. Find the distribution of $Y = -\ln(X)$ where $\ln$.

Solution:
Based on the transformation of random variables, we know that for univariate random variables

$$f_Y(y) = \frac{f_X(g^{-1}(y)}{|g'(g^{-1}(y))|}$$

where $g(x)$ is differentiable and monotonically increasing/decreasing function. In this case, $g(x) = -\ln(X)$ which is differentiable and monotonically decreasing function. Therefore, we can write the distribution of Y as

$$f_Y(y) = \frac{f_X(e^{-y})}{1/e^{-y}} = e^{-y} \quad \forall y > 0.$$

$f_Y(y)$ is an exponential distribution with mean 1.

b) Let X and Y be independent univariate $\mathcal{N}(0, 1)$ random variables. Let R denote the length of the vector $[X, Y]'$, and let Θ denote the angle the vector makes with the x-axis. In other words, if X and Y are the Cartesian coordinates of a random point in the plane, then $R \geq 0$ and $-\pi < \Theta \leq \pi$ are the corresponding polar coordinates. Find the joint density of R and Θ.

Solution: The transformation $[r, \theta]^T = G(x, y)$ is given by

$$r = \sqrt{x^2 + y^2},$$
$$\theta = \text{angle}(x, y).$$

The inverse transform $[x, y]^T = H(r, \theta)$ is the mapping that takes polar coordinates into the Cartesian coordinates. Hence, $H(r, \theta)$ is given by

$$x = r\cos\theta,$$
$$y = r\sin\theta.$$

The matrix $dH(r, \theta)$ is given by

$$dH(r, \theta) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix},$$

and $\det(dH(r, \theta) = r$. Then,

$$\begin{aligned} f_{R,\Theta}(r, \theta) &= f_{XY}(x, y)|x = r\cos\theta, y = r\sin\theta.|\det(H(r, \theta))|, \\ &= f_{XY}(r\cos\theta, r\sin\theta)r. \end{aligned}$$

Now, since X and Y are independent $\mathcal{N}(0,1)$, $f_{XY}(x,y) = f_X(x)f_Y(y) = e^{-(x^2+y^2)/2}/(2\pi)$, and

$$f_{R,\Theta}(r,\theta) = re^{-r^2/2}.\frac{1}{2\pi}, r \geq 0, -\pi < 0 \leq \pi.$$

Thus, R and Θ are independent, with R having a Rayleigh density and Θ having a uniform density.

(2) **Probability**

a) Let $X \sim \text{Unif}[0,1]$ be a uniformly-distributed random variable. Suppose we know $X+Y=1$ in advance, then

i) derive the distribution of Y.

ii) derive the distribution of $Z = \max\{X, Y\}$.

iii) compute the expectation of Z and $M = \min\{X, Y\}$.

Solutions:

i) For $y \in [0,1]$, we have

$$P(Y \leq y) = P(1-X \leq y) = P(X \geq 1-y) = 1-(1-y) = y.$$

Thus, $Y \sim \text{Unif}[0,1]$.

ii)

$$P(Z \leq z) = P(X \leq z, Y \leq z) = P(X \leq z, 1-X \leq z) = P(X \leq z, X \geq 1-z)$$

For $z \in [1/2, 1]$, we have $P(Z \leq z) = P(X \leq z, X \geq 1-z) = P(1-z \leq X \leq z) = 2z-1$
For $z \in [0, 1/2]$, $P(Z \leq z) = 0$.
We can see that the probability density of Z, i.e., $f_Z(z) = 2$, when $z \in [1/2, 1]$. Thus, $Z \sim \text{Unif}[1/2, 1]$.

iii) $E[Z] = 3/4$ because $Z \sim \text{Unif}[1/2, 1]$.
Since $E[M+Z] = E[X+Y]$, thus we have $E[M] = E[X+Y] - E[Z] = 1 - 3/4 = 1/4$.

(3) **Estimator performance**

a) Let $X_1, \ldots, X_N$ be N independent and identically distributed (i.i.d.) samples drawn from $\mathcal{N}(\mu, \sigma^2)$, where the mean μ is known in advance, while the variance σ^2 is unknown. We have three estimators to estimate the variance,

$$\hat{\sigma}_1^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2,$$

$$\hat{\sigma}_2^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \frac{1}{N}\sum_{j=1}^{N}X_j\right)^2,$$

$$\hat{\sigma}_3^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \frac{1}{N}\sum_{j=1}^{N}X_j\right)^2.$$

i) Check if $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, and $\hat{\sigma}_3^2$ are unbiased.

ii) For any biased estimator above, check if it is asymptotically unbiased.

Solution:

i)

$$E(\hat{\sigma}_1^2) = \frac{1}{N}\sum_{i=1}^{N} E\left(X_i - \mu\right)^2 = \frac{1}{N} * N\sigma^2 = \sigma^2,$$

$$\begin{aligned}
E(\hat{\sigma}_2^2) &= \frac{1}{N}\sum_{i=1}^{N} E\left(X_i - \frac{1}{N}\sum_{j=1}^{N} X_j\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N} E\left(\frac{N-1}{N}(X_i - \mu) - \frac{1}{N}\sum_{j\neq i}(X_j - \mu)\right)^2 \\
&\vdots \\
&= \frac{N-1}{N}\sigma^2 \\
E(\hat{\sigma}_3^2) &= \frac{N}{N-1}E(\hat{\sigma}_2^2) = \sigma^2.
\end{aligned}$$

Thus, $\hat{\sigma}_1^2$ and $\hat{\sigma}_3^2$ are unbiased estimators for σ^2, while $\hat{\sigma}_2^2$ is not.

ii) When $N \to +\infty$, we can see that $\hat{\sigma}_2^2 \to \sigma^2$, which means that $\hat{\sigma}_2^2$ is asymptotically unbiased.

(4) **Sufficient Statistics**

Find a sufficient statistic for the following distributions.

a) (Normal distribution)
We consider a joint normal distribution for which the mean μ is unknown, but the variance σ^2 is known:

$$f(x_1, \ldots, x_n|\mu) = (2\pi)^{-n/2}\sigma^{-n}\exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2)$$

what about the case when neither σ nor μ is known?

b) (Uniform distribution)
Now suppose the X_is are uniformly distributed on $[0, \theta]$ where θ is unknown, with the joint density given as

$$f(x_1, \ldots, x_n|\theta) = \theta^{-n}\mathbb{I}(x_i \leq \theta, \forall i)$$

where $\mathbb{I}(\cdot)$ is the indicator function.

c) (Gamma distribution)
Now suppose X_is have gamma distribution with β known and α unknown:

$$f(x_1, \ldots, x_n|\alpha) = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n}(\Pi_{i=1}^{n} x_i^{\alpha-1})\exp(-\beta\sum_{i=1}^{n} x_i)$$

Solution:

a)

$$\begin{aligned}
f(x_1, \ldots, x_n|\mu) &= (2\pi)^{-n/2}\sigma^{-n}\exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2) \\
&\vdots \\
&= (2\pi)^{-n/2}\sigma^{-n}\exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 - \frac{n\mu^2}{2\sigma^2}) \cdot \exp(\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i)
\end{aligned}$$

Since σ^2 is known,by the factorization theorem, we have the sufficient statistics as $T = \sum_{i=1}^{n} X_i$. When we don't know σ, by the factorization, we need to have jointly sufficient statistics given as

$$T_1 = \sum_{i=1}^{n} X_i, T_2 = \sum_{i=1}^{n} X_i^2$$

b) Note that $x_i \leq \theta, \forall i$ if and only if $\max\{x_1, x_2, \dots, x_n\} \leq \theta$. Thus we have

$$f(x_1, \dots, x_n|\theta) = \theta^{-n}\mathbb{I}(\max\{x_1, \dots, x_n\} \leq \theta)$$

And by the factorization theorem, this shows that $T = \max\{X_i\}$ is a sufficient statistic.

c) We can write

$$\prod_{i=1}^{n} x_i^{\alpha-1} = \exp((\alpha - 1)\sum_{i=1}^{n} \ln(x_i))$$

Thus by factorization, $T = \sum_{i=1}^{n} \ln(X_i)$ is a sufficient statistics.

(5) **Matlab Problem: Exploring Bias**. Let $x[n]$, $n = 0, ..., N-1$ be i.i.d. samples from a Normal distribution with mean μ and variance σ^2. In an algebraic mishap, you are given estimators of the form

$$\hat{\mu} = \frac{1}{N-1}\sum_{n=0}^{N-1} x[n]$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{n=0}^{N-1}(x[n] - \frac{1}{N}\sum_{n=0}^{n-1} x[n])^2$$

(i) **(5 pts)** Compute the theoretical bias of $\hat{\mu}$. Note that you have already computed the bias of the variance estimator in a previous problem.
Solution:
$E(\hat{\mu}) = \frac{1}{N-1}\mathbb{E}(\Sigma_{n=0}^{N-1} x[n] = 1/(N-1)\mu$. Then the bias is $1/(N-1) - 1)\mu$.

(ii) **(5 pts)** For a fixed variance, $\sigma^2 = 1$, vary μ from $0, 10, 20, 30, ..., 100$. Generate 1000 random samples x[n] of length N=50 in MATLAB and compute the estimator of μ for each realization. Compute the average value of the estimator, and create a table comparing the true value of μ and the bias of $\hat{\mu}$.
Solution: See HW1.m on canvas.

(iii) **(5 pts)** For a fixed mean $\mu = 0$, vary σ^2 from $1, 5, 10, 15, ..., 50$. Again generate 1000 random samples of length N=50 and compute the estimator of σ^2 for each realization. Compute the average value and variacne of the estimator and create a table comparing the true value of σ^2, the average estimate of σ^2, and the estimator variance for each value of σ^2.

(iv) **(5 pts)** Fix $\mu = 10$ and $\sigma^2 = 5$. Generate 1000 random samples for each N from $10, 50, 100, 200, ..., 1000$, and compute the average and variance of both estimators. What happens to the estimator bias and variance as N increases? Does the variance approach the Cramer Rao bound? **Solution:** As N increases, the variance of the estimators does approach the CRB.