Please provide complete and well-written solutions to the following exercises.

Due March 22, at the beginning of class.

Homework 5

Exercise 1 (Conditional Expectation as a Random Variable)**.** Let $X, Y, Z \colon \Omega \to \mathbf{R}$ be discrete or continuous random variables. Let A be the range of Y. Define $g\colon A \to \mathbf{R}$ by $g(y) := \mathbf{E}(X|Y = y)$, for any $y \in A$. We then define the **conditional expectation** of X given Y, denoted $\mathbf{E}(X|Y)$, to be the random variable $g(Y)$.

(i) Let X, Y be random variables such that (X, Y) is uniformly distributed on the triangle $\{(x, y) \in \mathbf{R}^2 \colon x \geq 0, y \geq 0, x + y \leq 1\}$. Show that

$$\mathbf{E}(X|Y) = \frac{1}{2}(1 - Y).$$

You only need to prove the following things for discrete random variables, or for continuous random variables (your choice).

(ii) Prove the following version of the Total Expectation Theorem

$$\mathbf{E}(\mathbf{E}(X|Y)) = \mathbf{E}(X).$$

- If X is a random variable, and if $f(t) := \mathbf{E}(X - t)^2$, $t \in \mathbf{R}$, then the function $f\colon \mathbf{R} \to \mathbf{R}$ is uniquely minimized when $t = \mathbf{E}X$. A similar minimizing property holds for conditional expectation. Let $h\colon \mathbf{R} \to \mathbf{R}$. Show that the quantity $\mathbf{E}(X - h(Y))^2$ is minimized among all functions $h\colon \mathbf{R} \to \mathbf{R}$ when $h(Y) = \mathbf{E}(X|Y)$. (Hint: use the previous item and (iii).)

(iii) Show the following:

$$\mathbf{E}(Xh(Y)|Y) = h(Y)\mathbf{E}(X|Y).$$

$$\mathbf{E}([\mathbf{E}(X|h(Y))]\,|Y) = \mathbf{E}(X|h(Y)).$$

(iv) Show the following

$$\mathbf{E}(X|X) = X.$$

$$\mathbf{E}(X + Y|Z) = \mathbf{E}(X|Z) + \mathbf{E}(Y|Z).$$

(v) If Z is independent of X and Y, show that

$$\mathbf{E}(X|Y, Z) = \mathbf{E}(X|Y).$$

(Here $\mathbf{E}(X|Y, Z)$ is notation for $\mathbf{E}(X|(Y, Z))$ where (Y, Z) is interpreted as a random vector, so that X is conditioned on the random vector (Y, Z).)

Exercise 2 (Conditional Jensen Inequality)**.** Prove Jensen's inequality for the conditional expectation. Let $X, Y \colon \Omega \to \mathbf{R}$ be random variables that are either both discrete or both continuous. Let $\phi\colon \mathbf{R} \to \mathbf{R}$ be convex. Then

$$\phi(\mathbf{E}(X|Y)) \leq \mathbf{E}(\phi(X)|Y)$$

If ϕ is strictly convex, then equality holds only if X is constant on any set where Y is constant. That is, (by an Exercise from the previous homework) equality holds only if X is a function of Y.

(Hint: first show that if $X \geq Z$ then $\mathbf{E}(X|Y) \geq \mathbf{E}(Z|Y)$.)

Exercise 3. Let Y, Z be a statistics, and suppose Z is sufficient for $\{f_\theta\colon \theta \in \Theta\}$. Show that $W := \mathbf{E}_\theta(Y|Z)$ does not depend on θ. That is, there is a function $t\colon \mathbf{R}^n \to \mathbf{R}$ that does not depend on θ such that $W = t(X)$, where X is the random sample.

Exercise 4. Let $X_1, \ldots, X_n$ be a random sample of size n, so that X_1 is a sample from the uniform distribution on the interval $[\theta - 1/2, \theta + 1/2]$, where $\theta \in \mathbf{R}$ is unknown.

- Show that $(X_{(1)}, X_{(n)})$ is minimal sufficient but not complete.
- The sample mean $\overline{X}$ might seem to be a reasonable estimator for θ, but it is not a function of the minimal sufficient statistic, so maybe it is not so good. Find an unbiased estimator for θ with smaller variance than $\overline{X}$ (for all θ). Then, examine the ratio of the variances (i.e. relative efficiency) for $\overline{X}$ and your estimator. (Don't try to find a UMVU; it does not exist! We will show this on the next homework.)

Exercise 5. Let $X_1, \ldots, X_n$ be a random sample of size $n = 2$, so that X_1 is a sample from exponential distribution with unknown parameter $\theta > 0$, so that X_1 has density $\theta e^{-x\theta} 1_{x>0}$.

Suppose we want to estimate the mean

$$g(\theta) := 1/\theta.$$

- Using the Rao-Blackwell Theorem (or any other method), find the UMVU for $g(\theta)$.
- Show that $\sqrt{X_1 X_2}$ has smaller mean squared error than the UMVU.
- Find an estimator with even smaller mean squared error, for all $\theta \in \Theta$.