CONDITIONAL EXPECTATIONS

1. This elementary problem is helpful in developing intuition about the concept of conditional expectation. Consider a probability space \((\Omega, \mathcal{F}, P)\) and two random variables on \(\Omega\) such that \(X\) takes finitely many distinct values \(x_1, \ldots, x_m\) and \(Z\)—finitely many distinct values \(z_1, \ldots, z_n\). As known from elementary probability theory the conditional probabilities are given by

\[
P[X = x_i | Z = z_j] = \frac{P[X = x_i; Z = z_j]}{P[Z = z_j]}
\]

and the conditional expectation of \(X\) given a value of \(Z\) equals

\[
E[X | Z = z_j] = \sum_{i=1}^{n} x_i P[X = x_i | Z = z_j]
\]

We can think of this expression as a function \(Y\) of \(\omega\), defined as \(Y(\omega) = E[X | Z = z_j]\) if \(Z(\omega) = z_j\). That is, \(Y\) is constant on the sets on which \(Z\) is constant.

Describe the \(\sigma\)-algebra \(\sigma(Z)\) generated by \(Z\) and prove that \(Y\) is measurable with respect to this \(\sigma\)-algebra. Prove that \(Y\) satisfies the definition of a conditional expectation of \(X\) with respect to \(\sigma(Z)\).

2. Here is another situation where it is easy to describe a conditional expectation explicitly. Suppose \(X\) and \(Z\) are random variables (on some probability space) with the joint density \(f_{X,Z}(x,z)\). Define

\[
f_Z(z) = \int_{\mathbb{R}} f_{X,Z}(x,z) \, dx.
\]

Prove that this is a Borel-measurable function of \(z\) which represents the density of the distribution of \(Z\). Define conditional density of \(X\) given \(Z\) as

\[
f_{X|Z}(x|z) = \frac{f_{X,Z}(x,z)}{f_Z(z)}
\]

if \(f_Z(z) \neq 0\) and as zero otherwise.

Now let \(h\) be a Borel function on \(\mathbb{R}\) such that

\[
E[|h(X)|] < \infty.
\]

Prove that the function

\[
g(z) = \int_{\mathbb{R}} h(x)f_{X|Z}(x|z) \, dx
\]

is a (version of) the conditional expectation \(E[h(X)|\sigma(Z)]\). Note that this is a Borel function of \(z\). This should not be surprising and the next problem shows that it is always the case.

3. Let \(\xi_1, \ldots, \xi_n\) be random variables on a measurable space \((\Omega, \mathcal{F})\). Prove that if \(\eta\) is a random variable measurable with respect to the \(\sigma\)-algebra \(\sigma(\xi_1, \ldots, \xi_n)\), then there exists a Borel function \(g : \mathbb{R}^n \to \mathbb{R}\) such that \(\eta = g(\xi_1, \ldots, \xi_n)\).

4. Existence and uniqueness of conditional expectations:

a) Let \(X\) be an integrable random variable on \((\Omega, \mathcal{F}, P)\) and \(\mathcal{G}\) a sub-\(\sigma\)-algebra of \(\mathcal{F}\). Prove that if \(Y_1\) and \(Y_2\) satisfy the requirements of the definition of \(E[X|\mathcal{G}]\) than \(Y_1 = Y_2\) \(P\)-almost surely.
Now assume that $X$ is square-integrable: $E[X^2] < \infty$. This implies integrability, because $P$ is a finite measure. Consider the Hilbert space $\mathcal{H} = L^2(\Omega, \mathcal{F}, P)$ and its closed subspace $\mathcal{K} = L^2(\Omega, \mathcal{G}, P)$. Prove that the orthogonal projection of $X$ onto $\mathcal{K}$ satisfies the definition of a conditional expectation $E[X|\mathcal{G}]$.

c) To extend the existence of conditional expectations to arbitrary $L^1$ random variables we need a simple Lemma: if $U$ is a bounded, nonnegative random variable, then its conditional expectation with respect to $\mathcal{G}$ (which exists by part b)) is also nonnegative.

d) Prove existence of a conditional expectation $E[X|\mathcal{G}]$ for any integrable $X$. Suggestion: it is enough to consider $X \geq 0$. Approximate $X$ from below by an increasing sequence of bounded random variables and use parts b) and c).

5. Fundamental properties of conditional expectations

Prove the following statements, without consulting books or notes, if possible. In what follows $E[|X|] < \infty$; $\mathcal{G}$ and $\mathcal{H}$ are sub-$\sigma$-algebras of $\mathcal{F}$.

a) if $Y = E[X|\mathcal{G}]$, then $E[X] = E[Y]$.

b) if $X$ is $\mathcal{G}$-measurable, then $E[X|\mathcal{G}] = X$ $P$-almost surely.

c) $E[a_1X_1 + a_2X_2|\mathcal{G}] = a_1E[X_1|\mathcal{G}] + a_2E[X_2|\mathcal{G}]$. Since conditional expectation is defined up to almost sure equivalence, a more precise statement would be: if $Y_i$ is a version of $E[X_i|\mathcal{G}]$ for $i = 1, 2$, then $a_1Y_1 + a_2Y_2$ is a version of $E[a_1X_1 + a_2X_2|\mathcal{G}]$. Make sure that you understand this last remark so well that in the future we do not need this level of precision.

d) (Conditional Jensen’s inequality—slightly harder.) If $c : \mathbb{R} \to \mathbb{R}$ is convex and $E[|c(X)|] < \infty$, then

$$E[c(X)|\mathcal{G}] \geq c(E[X|\mathcal{G}])$$

$P$-almost surely.

e) If $\mathcal{H}$ is a sub-$\sigma$-algebra of $\mathcal{G}$, then

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}]$$

f) If $Z$ is $\mathcal{G}$-measurable and bounded then

$$E[ZX|\mathcal{G}] = ZE[X|\mathcal{G}]$$

The same is true if we assume that $Z \in L^p(\Omega)$ and $X \in L^q(\Omega)$ with $\frac{1}{p} + \frac{1}{q} = 1$.

g) For this problem review the definition of independent $\sigma$-algebras. If $\mathcal{H}$ is independent of $\sigma(\sigma(X), \mathcal{G})$ (the $\sigma$-algebra generated by $\mathcal{G}$ and $X$), then

$$E[X|\sigma(\mathcal{G}, \mathcal{H})] = E[X|\mathcal{G}]$$

What does the special case when $\mathcal{G} = \{\emptyset, \Omega\}$ and $X$ is independent of $\mathcal{H}$ say?