CONDITIONAL EXPECTATIONS

The conditional expectation $\mathbb{E}^{\mu}[X|\mathcal{G}]$ of X given $\mathcal{G} \subset \mathcal{F}$ on a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$ is defined in terms of Radon-Nikodym's theorem. More precisely, if X is non-negative or X is σ-integrable on $\mathcal{G}$, that is, there is a sequence of $G_n \in \mathcal{G}$ such that $\mathbb{E}^{\mu}[X : G_n] < \infty$ for each n and $\bigcup_{n=1}^{\infty} G_n = \Omega$.

1) *Definition of Conditional expectations.* Suppose $(\Omega, \mathcal{F}, \mu)$ is a measure space, $\mathcal{G} \subseteq \mathcal{F}$ is a sub-algebra and μ is σ-finite on $\mathcal{G}$. If X is a non-negative real random variable which is σ-integrable on $\mathcal{G}$, then there is a $\mathcal{G}$-measurable random variable $\mathbb{E}^{\mu}[X|\mathcal{G}]$, the conditional expectation of X, which is a unique (up to almost everywhere) function Y such that

1) Y is $\mathcal{G}$-measurable,

2) $\mathbb{E}^{\mu}[Y 1_A] = \mathbb{E}^{\mu}[X 1_A]$ for every $A \in \mathcal{G}$.

A random variable Y (either non-negative or integrable) which satisfies conditions 1) and 2) above is called the conditional expectation of a random variable X , denoted by $\mathbb{E}^{\mu}[X|\mathcal{G}]$.

Therefore, if a random variable X is non-negative and σ-integrable on $\mathcal{G}$, then its conditional expectation $\mathbb{E}^{\mu}[X|\mathcal{G}]$ exists and unique up to almost everywhere.

In what follows, let us work with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{G} \subseteq \mathcal{F}$ is a sub σ-algebra. The conditional expectation of X (if exists) will be denoted by $\mathbb{E}[X|\mathcal{G}]$.

2. Conditional expectations for integrable functions. Suppose X is integrable, thus X^+ and X^- are non-negative, $\mathcal{F}$-measurable and integrable, thus $\mathbb{E}[X^{\pm}|\mathcal{G}]$ are defined, $\mathcal{G}$-measurable, and integrable. Therefore both $\mathbb{E}[X^{\pm}|\mathcal{G}]$ are finite almost surely, so that

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}\left[X^+|\mathcal{G}\right] - \mathbb{E}\left[X^-|\mathcal{G}\right]$$

is integrable. $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and $\mathbb{E}[X : A] = \mathbb{E}[\mathbb{E}(X|\mathcal{G}) : A]$ for every $A \in \mathcal{G}$, so that $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation of X.

If X is $\mathcal{F}$-measurable and non-negative, then for each n, $X \wedge n$ is bounded and $X \wedge n \uparrow f$. Thus $\mathbb{E}[X \wedge n|\mathcal{G}]$ is defined for each n, and $\mathbb{E}[X \wedge n|\mathcal{G}]$ is increasing, its limit Y exists. Y is $\mathcal{G}$-measurable, and for every $A \in \mathcal{G}$, according to MCT, we have $\mathbb{E}[X : A] = \mathbb{E}[Y : A]$, so that Y is the conditional expectation of X, denoted by $\mathbb{E}[X|\mathcal{G}]$.

3. Example. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $A \in \mathcal{F}$ with $0 < \mathbb{P}(A) < 1$. Let $\mathcal{G} = \sigma(A)$. If $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ then

$$\mathbb{E}[X|\mathcal{G}] = \frac{\mathbb{E}[X : A]}{\mathbb{P}(A)} \mathbf{1}_A + \frac{\mathbb{E}[X : A^c]}{\mathbb{P}(A^c)} \mathbf{1}_{A^c}.$$

In general, if $\{A_j\}$ is a countable partition of Ω, i.e. $\cup_j A_j = \Omega$, $\{A_j\}$ are disjoint and $\mathbb{P}(A_j) > 0$, then

$$\mathbb{E}[X|\mathcal{G}] = \sum_{j=1}^{\infty} \frac{\mathbb{E}[X : A_j]}{\mathbb{P}(A_j)} \mathbf{1}_{A_j}$$

where $\mathcal{G} = \sigma\{A_j : j = 1, 2, \cdots\}$.

4. Notations. The following convention on conditional expectations will be assumed. If Z is a random variable, then the conditional expectation of X given Z, denoted by $\mathbb{E}[X|Z]$, is defined to be the conditional expectation of X given $\sigma(Z)$. If $Z_1, \cdots, Z_n$ is a finite family of random variables, then we define

$$\mathbb{E}[X|Z_1, \cdots, Z_n] = \mathbb{E}[X|\sigma(Z_1, \cdots, Z_n)].$$

In general, if $\{Z_\alpha\}_{\alpha\in\Lambda}$ is a family of random variables, then

$$\mathbb{E}[X|Z_\alpha; \alpha \in \Lambda] = \mathbb{E}[X|\sigma(\{Z_\alpha\}_{\alpha\in\Lambda})] .$$

5. Example. Let X and Z be two random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with continuous joint probability density function $p(x, z)$, i.e.

$$\mathbb{P}\{(X, Z) \in D\} = \iint_D p(x, z)dxdz.$$

Then

$$\mathbb{E}[f(X)|Z] = \frac{\int_{\mathbb{R}} f(x)p(x, Z)dx}{\int_{\mathbb{R}} p(x, Z)dx}$$

where f is Borel measurable, non-negative or/and $f(X)$ is integrable. In fact, formally

$$\mathbb{P}[X = x|Z = z] = \frac{\mathbb{P}(X = x, Z = z)}{\mathbb{P}(Z = z)} = \frac{p(x, z)}{\int_{\mathbb{R}} p(x, z)dx}.$$

6. Properties of the conditional expectations.

6.1) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}(X)$, i.e. the expectation of conditional expectation doesn't change. If X is integrable, and X is $\mathcal{G}$-measurable, then $\mathbb{E}[X|\mathcal{G}] = X$. If Z is $\mathcal{G}$-measurable, then$\mathbb{E}[ZX|\mathcal{G}] = Z\mathbb{E}[X|\mathcal{G}]$.

6.2) $X \to \mathbb{E}(X|\mathcal{G})$ is linear, additive and positive.

6.3) Convergence Theorems. 6.3.1) MCT for conditional expectations: If $0 \le X_n \uparrow X$ then $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$. *8.3.2) Fatou's Lemma:* If $X_n \ge 0$, then $\mathbb{E}[\liminf X_n|\mathcal{G}] \le \liminf \mathbb{E}[X_n|\mathcal{G})]$. *8.3.3) Dominated Convergence*: If $|X_n| \le Z$ for some $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\lim X_n = X$, then $\mathbb{E}[X_n|\mathcal{G}] \Rightarrow \mathbb{E}[X|\mathcal{G}]$.

6.4) If $\mathcal{G}_2 \subset \mathcal{G}_1 \subset \mathcal{F}$, then $\mathbb{E}\{\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2\} = \mathbb{E}[X|\mathcal{G}_2]$ (this is called the power law for conditional expectations).

7. Jensen's inequality for conditional expectations. If φ is convex, and both X and $\varphi(X)$ are integrable, then

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \le \mathbb{E}[\varphi(X)|\mathcal{G}]$$

almost surely.

Let us prove the Jensen inequality. Recall that φ is convex on $\mathbb{R}$ if

$$\varphi(\lambda s + (1-\lambda)t) \leq \lambda\varphi(s) + (1-\lambda)\varphi(t)$$

for all $s, t \in \mathbb{R}$ and $\lambda \in [0,1]$, which is equivalent to that

$$\frac{\varphi(u) - \varphi(s)}{u - s} \leq \frac{\varphi(t) - \varphi(u)}{t - u}$$

for any $s < u < t$ (with $u = \lambda s + (1-\lambda)t$). In particular, the right-derivative

$$\varphi'_+(s) = \lim_{t \downarrow s} \frac{\varphi(t) - \varphi(s)}{t - s} = \inf_{t > s} \frac{\varphi(t) - \varphi(s)}{t - s}$$

exists. Similarly

$$\varphi'_-(t) = \lim_{s \uparrow t} \frac{\varphi(t) - \varphi(s)}{t - s} = \sup_{s < t} \frac{\varphi(t) - \varphi(s)}{t - s}.$$

and both $t \to \varphi'_\pm(t)$ are increasing. By definition, for $s < t$ we have

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \varphi'_-(t)$$

that is

$$\varphi(s) \geq \varphi(t) + \varphi'_-(t)(s - t)$$

for $s < t$. While if $s > t$, then

$$\frac{\varphi(s) - \varphi(t)}{s - t} \geq \varphi'_+(t) \geq \varphi'_-(t)$$

we thus also have

$$\varphi(s) \geq \varphi(t) + \varphi'_-(t)(s - t).$$

Therefore, for a convex function φ, we have

$$\varphi(s) \geq \varphi(t) + \varphi'_-(t)(s - t) \text{ for all } s. \tag{7.1}$$

Applying (7.1) $t = \mathbb{E}[X|\mathcal{G}]$ and $s = X$, to obtain

$$\varphi(X) \geq \varphi(\mathbb{E}[X|\mathcal{G}]) + \varphi'_-(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}]).$$

Now $t \to \varphi'_-(t)$ is increasing, so that it is Borel measurable, thus $\varphi'_-(\mathbb{E}[X|\mathcal{G}])$ is $\mathcal{G}$-measurable. Taking conditional expectation we deduce that

$$\begin{aligned}
\mathbb{E}[\varphi(X)|\mathcal{G}] &\geq \varphi(\mathbb{E}[X|\mathcal{G}]) + \mathbb{E}\left[\varphi'_-(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}])|\mathcal{G}\right] \\
&= \varphi(\mathbb{E}[X|\mathcal{G}]) + \varphi'_-(\mathbb{E}[X|\mathcal{G}])\,\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])|\mathcal{G}] \\
&= \varphi(\mathbb{E}[X|\mathcal{G}]).
\end{aligned}$$