§ 3.3 The Projection Property of Conditional Expectation

Let $\mathcal{F}$ be a $\sigma$-field and $L^2(\mathcal{F})$ is the collection of random variables $Z$ on $\Omega$, satisfying

1. $Z$ has finite second moment: $EZ^2 < \infty$,

2. The information carried by $Z$ is contained in $\mathcal{F}$: $\sigma(Z) \subseteq \mathcal{F}$. If $\mathcal{F} = \sigma(Y)$, this means that $Z$ is a function of $Y$.

**The Projection Property**

Let $X$ be a random variable with $EX^2 < \infty$. The conditional expectation $E(X \mid \mathcal{F})$ is that random variable in $L^2(\mathcal{F})$ which is closest to $X$ in the mean square sense. This means that

$$E[X - E(X \mid \mathcal{F})]^2 = \min_{Z \in L^2(\mathcal{F})} E(X - Z)^2.$$  \hspace{1cm} (1.39)

In this sense, $E(X \mid \mathcal{F})$ is the *projection of the random variable $X$ on the space* $L^2(\mathcal{F})$ of the random variables $Z$ carrying part of the information $\mathcal{F}$; also see the comments to Figure 1.4.16.

If $\mathcal{F} = \sigma(Y)$, $E(X \mid Y)$ is that function of $Y$ which has a finite second moment and which is closest to $X$ in the mean square sense.
Figure 1.4.16 An illustration of the projection property of the conditional expectation $E(X|F)$. We mention that $<Z,Y> = E(ZY)$ for $Z$, $Y$ with $EZ^2 < \infty$ and $EY^2 < \infty$ defines an inner product and $\|Z-Y\| = \sqrt{<Z-Y,Z-Y>}$ a distance between $Z$ and $Y$. As in Euclidean space, we say that $Z$ and $Y$ are orthogonal if $<Z,Y> = 0$. In this sense, $E(X|F)$ is the orthogonal projection of $X$ on $L^2(F)$: $<X - E(X|F), Z> = 0$ for all $Z \in L^2(F)$, and $<X - Z, X - Z>$ is minimal for $Z = E(X|F)$.

Since we have shown

$$E(B_t | B_x, x \leq s) = B_s$$

and

$$E(B_t^2 - t | B_x, x \leq s) = B_s^2 - s$$

We may say that the best predictors of the future values $B_t$ and $B_t^2 - t$, given the information about the Brownian motion until the present times, are $B_s$ and $B_s^2 - s$. 
Sketch of Proof of the Projection Property

Let \( Z \) be any r.v. in \( L^2(\mathcal{F}) \). Then

\[
E(X-Z)^2 = E[(X-Z') + (Z'-Z)]^2
= E[(X-Z')^2] + E[(Z'-Z)^2] + 2E[(X-Z')(Z'-Z)] \quad \ldots (\ast)
\]

where \( Z' = E[X|\mathcal{F}] \).

Applying Rule 5,

\[
E[(X-Z')(Z'-Z)|\mathcal{F}] = (Z'-Z)E[(X-Z')|\mathcal{F}]
\]

By Rule 1 and 4

\[
E[X-Z'|\mathcal{F}] = E[X|\mathcal{F}] - E[Z'|\mathcal{F}]
= Z' - Z' = 0
\]

So the last term in (\ast) is zero. Thus we have proved

\[
E(X-Z)^2 = E(X-Z')^2 + E(Z-Z')^2
\]

Hence

\[
E(X-Z)^2 \geq E(X-Z')
\]

for all r.v. \( Z \) in \( L^2(\mathcal{F}) \).