Notes on the Conditional Expectation

Cheng Wang¹, a, Xueyi Liu², Hailei Zou³

¹,²,³Department of Applied Mathematics, China Jiliang University, Hangzhou, 310018, China

aemail: wangc628@126.com

Keywords: Conditional expectation, Regular conditional distribution function, Independence, Yule process.

Abstract. We first present a formula on the conditional expectation by the regular conditional distribution function. Using this formula, we can obtain a corollary under the condition of independence which applies in many cases. Then some examples are given to illustrate the applications of the results.

Introduction

Given a random variable $X$ on a probability space $(\Omega, F, P)$, and a sub-$\sigma$-field $G \subset F$, the conditional expectation of $X$ given $G$ denoted by $E[X | G]$ is well known and is an important mark of modern probability. Kallenberg (¹) listed its precise definitions and many properties. Ikeda(²) and Jiagang Wang(³) discuss the conditional expectation in terms of regular conditional probability and regular conditional distribution function respectively.

This paper uses the regular conditional distribution function respectively to give a formula on the conditional expectation when some of the random variables are $G$-measurable. This result is intuitive and used frequently. The calculations of many typical examples(⁴) potentially use this property, but no authors give detailful explanations.

In the following, we let $(\Omega, F, P)$ be a probability space, $G$ be a sub-$\sigma$-field of $F$, $X_1, X_2, \ldots, X_n$ be random variables on $(\Omega, F, P)$.

Definition 1. (³) A function $F(x_1, x_2, \ldots, x_n, \omega)$ on $\mathbb{R}^n \times \Omega$ is called a regular conditional distribution function of $(X_1, X_2, \ldots, X_n)$ given $G$, if it satisfies all the following:

1. $F(x_1, x_2, \ldots, x_n, \omega)$ is $G$-measurable for fixed $x_1, x_2, \ldots, x_n$;
2. $F(x_1, x_2, \ldots, x_n, \omega)$ is an $n$-dimensional distribution function for fixed $\omega$;
3. $F(x_1, x_2, \ldots, x_n, \omega) = E[I_{(-\infty, x_1)} \cdot (-\infty, x_2) \ldots (-\infty, x_n)(X_1, \ldots, X_n) | G]$, a.s.

We denote this regular conditional distribution function by $F_G(x_1, x_2, \ldots, x_n, \omega)$.

Lemma 1. (³) Suppose $f(x_1, x_2, \ldots, x_n)$ is a Borel function such that $f(X_1, X_2, \ldots, X_n)$ is integrable. Then we have

$$E[f(X_1, \ldots, X_n) | G] = \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dF_G(x_1, x_2, \ldots, x_n, \omega) \text{ a.s.}$$

Lemma 2. (²) Let $X, Y$ be two integrable random variables on $(\Omega, F, P)$, $X$ be $G$-measurable, and $XY$ be integrable. Then

$$E[XY | G] =XE[Y | G] \quad \text{a.s.}$$

Main Results

Theorem 1 Let $X, Y$ be two random variables, $X$ be $G$-measurable, and $f$ be a Borel function defined on $\mathbb{R}^2$ such that $f(X, Y)$ is integrable. Then

$$E[f(X, Y) | G] = \int f(X(\omega), y) dF_G(y, \omega) \text{ a.s.} \quad (1)$$
where $F_G(y)$ is the regular conditional distribution function of $Y$ given $G$.

Proof. We first prove (1) for all $f = 1_{A \times B}(x, y)$, where $A$ and $B$ are any Borel sets on $\mathbb{R}$. By lemma 1 and lemma 2, we have

$$E[1_{A \times B}(X, Y) | G] = 1_A(X(\omega))E[1_B(Y) | G]$$

$$= 1_A(X(\omega))\int_B(y) dF_G(y, \omega) = \int_A(X(\omega))1_B(y) dF_G(y, \omega)$$

$$= \int_{A \times B}(X(\omega), y) dF_G(y, \omega)$$

i.e., (1) holds for $f = 1_{A \times B}(x, y)$.

Let $M$ denote the set \( \{D: (1) \text{ holds for } f = 1_D(x, y)\} \), and $N$ denote the set \( \{A \times B: A, B \text{ are Borel sets}\} \). Obviously, $M$ is a $\lambda$-system and $N$ is a $\pi$-system. $M$ contains $N$ due to the above proof, so $M$ contains the $\sigma$-field generated by $N$. That is $M$ contains all Borel sets on $\mathbb{R}^2$.

Thus (1) holds for all $f = 1_D(x, y)$, where $D$ is any Borel set. Hence (1) holds for all simple functions on $\mathbb{R}^2$.

Furthermore, every nonnegative measurable function $f$ satisfies (1) by Levy monotone convergence theorem. As for a general Borel function $f(x, y)$, we write $f = f^+ - f^-$, where $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$. Since (1) holds for both $f^+$ and $f^-$, it obviously holds for $f$.

The proof is completed.

Remark 1. The well-known Lemma 2 is a special case of the above Theorem when we take $f = XY$. In fact, by Theorem 1,

$$E(XY | G) = \int X(\omega) y dF_G(y, \omega)$$

$$= X(\omega)\int y dF_G(y, \omega) = X(\omega)E[Y | G]$$

the last “=” is due to Lemma 1.

Recalling the independence between a random variable and a $\sigma$-field([1]), we have the following corollary of Theorem 1.

Corollary 1. Let $X, Y$ be two random variables, $X$ be $G$-measurable and $Y$ be independent of $G$. Then for any Borel function $f$ such that $f(X, Y)$ is integrable, we have

$$E[f(X, Y) | G] = E[f(x, Y)]_{x \sim X(\omega)}$$

Proof. Since $Y$ is independent of $G$, $F_G(y, \omega) = F_Y(y)$, where $F_Y(y)$ is the distribution of $Y$. So from Theorem 1,

$$E[f(X, Y) | G] = \int f(X(\omega), y) dF_G(y, \omega)$$

$$= \int f(X(\omega), y) dF_Y(y) = E[f(x, Y) | G]_{x \sim X(\omega)}.$$

Remark 2. The derivation of B-S formula for European options’ price([5], page 118) is implicitly uses this corollary.

Corollary 2. Suppose $X, Y$ are random variables, and $G$ is a $\sigma$-algebra. If $X$ is independent of $\sigma(Y, G)$ (the $\sigma$-algebra generated by $Y$ and $G$). Then we have

$$E[XY | G] = EX \cdot E[Y | G].$$

Proof. By the tower property of conditional expectation and corollary 1, we have
\[
E[XY \mid G] = E[E[XY \mid G] \mid \sigma(Y, G)] \\
= E[E[XY \mid \sigma(Y, G)] \mid G] \\
= E[E(X \mid Y) \mid \sigma(Y) \mid G] \\
= E[E(X \mid Y) \mid Y = y] \\
= E[X \mid Y = y] E[Y \mid G].
\]

The following Example 1 is an intuitive illustration of Theorem 1. Example 2 appears frequently in many textbooks but the authors always omit the detailful explanation on the crucial step in the calculation. Here Corollary 2 makes the crucial step more explicit.

Example 1. Let \( X \) be a discrete random variable with distribution \( P(X = 1) = 0.2, P(X = 2) = 0.3, P(X = 3) = 0.5 \). Suppose the regular conditional distribution function of \( Y \) given \( \sigma(X) \) (the \( \sigma \)-field generated by \( X \)) is

\[
F_{\sigma(X)}(y, \omega) = \begin{cases} 
1 - e^{-X(y)}y, & y > 0 \\
0, & y \leq 0.
\end{cases}
\]

Then

\[
E(e^{-X^2} \mid \sigma(X)) = \int e^{-X^2}dyF_{\sigma(X)}(y, \omega) = \int_0^{\infty} e^{-X^2}X(\omega)e^{-X(y)}dy
\]

\[
= X(\omega) \int_{(0, \infty)} e^{-(X^2+X(y))y}dy
\]

\[
= \begin{cases} 
1/2, & X(\omega) = 1, \\
1/3, & X(\omega) = 2, \\
1/4, & X(\omega) = 3.
\end{cases}
\]

Example 2. Suppose a cashflow comes in according to a Poisson process with intensity \( \lambda \). The coming cash each time \( (C_i) \) is often assumed to be a random variable with the same normal distributions and the cashflows are independent of the Poisson process. The discount rate is \( r \). Calculate the expected present value (PV) of all the cashes during the time interval \([0,t]\).

We all know the expectation is

\[
E[\sum_{i=1}^{N_t} C_i e^{-rt_i}] = E[E[\sum_{i=1}^{N_t} C_i e^{-rt_i} \mid N_t]]
\]

\[
= \sum_{n=1}^{\infty} E[\sum_{i=1}^{n} C_i e^{-rt_i} \mid N_t = n] P(N_t = n).
\]

The textbooks often straightly display

\[
E[\sum_{i=1}^{n} C_i e^{-rt_i} \mid N_t = n] = EC[\sum_{i=1}^{n} e^{-rt_i} \mid N_t = n]
\]

but never say why.

Now from Corollary 2, we know

\[
E[C_i e^{-rt_i} \mid N_t = n] = C_i EC[e^{-rt_i} \mid N_t = n],
\]

since \( C_i \) is independent of \( \sigma(\tau_i, N_t) \) (this is in that the cashflows are independent of the Poisson process). As far, we know why \( C_i \) can be drawn out of the expectation symbol.
Theorem 1 can easily be extended to the case of high dimensions as follows.

Theorem 2. Suppose $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are random variables on $(\Omega, F, P)$, and $X_1, X_2, \ldots, X_n$ are $G$-measurable. For every integrable $f(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m)$, we have

$$E[f(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m) | G] = \int f(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), y_1, y_2, \ldots, y_m) dF(y_1, y_2, \ldots, y_m, \omega)$$

Acknowledgement

In this paper, the research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LY14A010026 and LY14F030020.

References