Stat/219 Math 136 - Stochastic Processes

Notes on Section 4.1.2

1 Independence, Uncorrelated-ness, and Something in Between

Suppose (X, Y) is a random vector defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}(X^2) < \infty$, $\mathbb{E}(Y^2) < \infty$. The following implications are true.

$$\begin{array}{ccccc} & & \boldsymbol{\mathbb{E}(X|Y) = \mathbb{E}(X)} & & \\ & \nearrow & & \searrow & \\ \textbf{\textit{X} and \textit{Y} are independent} & & & & \textbf{\textit{X} and \textit{Y} are uncorrelated} \\ & \searrow & & \nearrow & \\ & & \boldsymbol{\mathbb{E}(Y|X) = \mathbb{E}(Y)} & & \end{array} \quad \textbf{(1\textit{A})}$$

For a proof of the first statements on the left above, assume X and Y are independent. Then the constant $\mathbb{E}(X)$ is measurable with respect to $\sigma(Y)$. Since X and $\sigma(Y)$ are independent, for any $G \in \sigma(Y)$ we have

$$\begin{aligned} \int_G \mathbb{E}(X) d\mathbb{P} = \mathbb{E}(X)\mathbb{E}(I_G) &= \mathbb{E}(XI_G) \text{ by independence} \\ &= \int_G X d\mathbb{P} \end{aligned}$$

Thus $\mathbb{E}(X|Y) = \mathbb{E}(X)$. Similarly we have $\mathbb{E}(Y|X) = \mathbb{E}(Y)$.

Next we prove the statements on the right side of (1A). Suppose that $\mathbb{E}(Y|X) = \mathbb{E}(Y)$. Using the tower property and then taking out what is known, we have $\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|X)) = \mathbb{E}(X(\mathbb{E}(Y|X)) = \mathbb{E}(X)\mathbb{E}(Y)$. Thus X and Y are uncorrelated. Similarly, $\mathbb{E}(X|Y) = \mathbb{E}(X)$ implies that X and Y are uncorrelated.

However, the converse implications are not true in general.

Counterexample 1: X **and** Y **uncorrelated does not imply** $\mathbb{E}(X|Y) = \mathbb{E}(X)$

Let $\Omega = \{-1, 0, 1\}$ with $\mathbf{P}(\{\omega\}) = 1/3$ for each $\omega \in \Omega$. Let $Y(\omega) = \omega$ and $X(\omega) = I_{\{0\}}(\omega)$. Then $XY = 0$ so $\mathbb{E}(XY) = 0$. Also, $\mathbb{E}(Y) = 0$, and so $\mathbb{E}(XY) = 0 = \mathbb{E}(X)\mathbb{E}(Y)$; that is, X and Y are uncorrelated. However, since X is measurable with respect to $\sigma(Y)$ we have $\mathbb{E}(X|Y) = X$ which is never equal to $\mathbb{E}(X) = 1/3$.

Counterexample 2: $\mathbb{E}(X|Y) = \mathbb{E}(X)$ **does not imply that** X **and** Y **are independent**

Consider the independent random variables X and S of Exercise 3.2.12. Then $\mathbb{E}(SX|X) = X\mathbb{E}(S|X) = X\mathbb{E}(S) = 0$. However, as shown in Exercise 3.2.12, X and SX are not independent.

Counterexample 3: $\mathbb{E}(X|Y) = \mathbb{E}(X)$ **does not imply that** $\mathbb{E}(Y|X) = \mathbb{E}(Y)$

Let $\Omega = \{-1, 0, 1\}$ with $\mathbb{P}(\{\omega\}) = 1/3$ for $\omega \in \Omega$. Consider $Y(\omega) = \omega$ and $X(\omega) = I_{\{0\}}(\omega)$. Then $\mathbb{E}(Y) = 0$ and $\mathbb{E}(Y|X) = 0 \times I_{\{0\}}(\omega) + (-1 \times 1/2 + 1 \times 1/2)I_{\{-1,1\}}(\omega) = 0$, so $\mathbb{E}(Y|X) = \mathbb{E}(Y)$. But $\mathbb{E}(X) = \mathbb{P}(\{0\}) = 1/3$ and since X is $\sigma(Y)$-measurable $\mathbb{E}(X|Y) = X = I_{\{0\}}(\omega)$, so for each $\omega \in \Omega$, $\mathbb{E}(X|Y)(\omega) \neq \mathbb{E}(X)$.

1.1 Relationships in terms of expected values

It is instructive to look at these relationships in (1A) in terms of equivalent statements about expectations of functions of the random variables:

$$
\begin{aligned}
X \text{ and } Y \text{ are independent} \quad &\Leftrightarrow \quad \mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)] \text{ for all bounded, measurable } g \text{ and } h \\
\mathbb{E}(X|Y) = \mathbb{E}(X) \quad &\Leftrightarrow \quad \mathbb{E}[Xh(Y)] = \mathbb{E}[X]\mathbb{E}[h(Y)] \text{ for all bounded, measurable } h \\
X \text{ and } Y \text{ are uncorrelated} \quad &\Leftrightarrow \quad \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}
$$

The first line is due to Proposition 1.4.41. The second line is due to the definition of conditional expectation in the L^2 case; see Definition 2.1.3 and equation (2.1.2).

In case X is square integrable it is possible to replace bounded g, h in the above statements with g, h that satisfy $\mathbb{E}[|g(X)|^2] < \infty$, $\mathbb{E}[|h(Y)|^2] < \infty$. From the right side of the lines above, it should be clear why independence is the strongest statement while uncorrelatedness is the weakest, and the second line gives a property strictly "in between." (It should also be apparent why there is an "asymmetry" in the "in between" property, as witnessed in Counterexample 3 above.)

1.2 Equivalence in the case of Gaussian Random Vectors

If (X, Y) is a Gaussian random vector then these notions are equivalent:

$$X \text{ and } Y \text{ are independent} \Leftrightarrow \mathbb{E}(X|Y) = \mathbb{E}(X) \text{ and } \mathbb{E}(Y|X) = \mathbb{E}(Y) \Leftrightarrow X \text{ and } Y \text{ are uncorrelated}$$

In order to show this, assume (X, Y) is a Gaussian random vector and X and Y are uncorrelated. Let $\mu = (\mu_X, \mu_Y)$ be the mean vector and $\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{pmatrix}$ be the covariance matrix. Since X and Y are uncorrelated, $\sigma_{X,Y} = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = 0$. So, $(\theta, \Sigma\theta) = \sigma_X^2\theta_1^2 + \sigma_Y^2\theta_2^2$ for $\theta = (\theta_1, \theta_2)$ and we have

$$
\begin{aligned}
\Phi_{X,Y}(\theta_1, \theta_2) &= \exp[-(1/2)(\sigma_X^2\theta_1^2 + \sigma_Y^2\theta_2^2) + i(\mu_X\theta_1 + \mu_Y\theta_2)] \\
&= \exp[-(1/2)\sigma_X^2\theta_1^2 + i\mu_X\theta_1]\exp[-(1/2)\sigma_Y^2\theta_2^2 + i\mu_Y\theta_2] = \Phi_X(\theta_1)\Phi_Y(\theta_2)
\end{aligned}
$$

where the last equality is due to the fact that any individual component of a Gaussian random vector is a Gaussian random variable. Thus, by Proposition 3.2.6 X and Y are independent, and the remaining equivalences then follow from (1A).

It is important to recall that the assumption that (X, Y) is a Gaussian random vector is stronger than just having X and Y be Gaussian random variables. For example see Exercise 3.2.12, in which we showed that SX and X are two uncorrelated Gaussian random variables, but (SX, X) is not a Gaussian random vector.

2 Applications to Square-Integrable Martingales

In this section we consider a stochastic process $\{X_n\}$ which is square integrable ($\mathbb{E}(X_n^2) < \infty$ for all n) and its canonical filtration $\{\mathcal{F}_n\}$ ($\mathcal{F}_n = \sigma(X_0, \dots, X_n)$). Let $D_0 = X_0$ and $D_n = X_n - X_{n-1}, n = 1, 2, \dots$ be the associated difference (or increment) sequence. Using Corollary 1.2.17 we can verify that $\mathcal{F}_n = \sigma(D_0, \dots, D_n)$.

It is easy to verify that if $\{D_n\}$ is a sequence of independent random variables with $\mathbb{E}(D_n) = 0$ then $\{(X_n, \mathcal{F}_n)\}$ is a martingale. However, independence of $\{D_n\}$ is not necessary in order for $\{(X_n, \mathcal{F}_n)\}$ to be a martingale, since as long as $\mathbb{E}(D_{n+1}|\mathcal{F}_n) = 0$ for all $n = 0, 1, 2, \dots$ we have

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}(X_n + D_{n+1}|\mathcal{F}_n) = X_n + \mathbb{E}(D_{n+1}|\mathcal{F}_n) = X_n.$$

Also, if $\{(X_n, \mathcal{F}_n)\}$ is a martingale, we must have $\mathbb{E}(D_{n+1}|\mathcal{F}_n) = 0$.

Note that the assumption $\mathbb{E}(D_{n+1}|\mathcal{F}_n) = 0$ is equivalent to assuming that both $\mathbb{E}(D_{n+1}) = 0$ and

$$\mathbb{E}(D_{n+1}|\mathcal{F}_n) = \mathbb{E}(D_{n+1}). \qquad (2A)$$

From Definition 2.1.3 and equation (2.1.2), (2A) is equivalent to: $\mathbb{E}[(D_{n+1} - \mathbb{E}(D_{n+1}))V]$ for any square integrable random variable V which is measurable with respect to $\mathcal{F}_n$. Recalling that $\mathcal{F}_n = \sigma(D_0, \dots, D_n)$, by Theorem 1.2.14 we have that for any such V there is a Borel measurable function $h : \mathbb{R}^{n+1} \mapsto \mathbb{R}$ for which $V = h(D_0, \dots, D_n)$. Thus we have that (2A) is equivalent to

$$\mathbb{E}[D_{n+1}h(D_0, \dots, D_n)] = \mathbb{E}[D_{n+1}]\mathbb{E}[h(D_0, \dots, D_n)] \text{ for all measurable } h \text{ with } \mathbb{E}[|h(D_0, \dots, D_n)|^2] < \infty.$$

Notice that this property is the same as the "in between" property mentioned in Section 1 above (for the two-dimensional case). We also see that this property is related to the orthogonality property discussed in Section 2.1. This motivates the following definition (Definition 4.1.2): We say that $D_n \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ is an *orthogonal sequence* if $\mathbb{E}[D_{n+1}h(D_0, D_1, \dots, D_n)] = \mathbb{E}[D_{n+1}]\mathbb{E}[h(D_0, D_1, \dots, D_n)]$ for any $n = 0, 1, 2, \dots$ and every Borel function $h : \mathbb{R}^{n+1} \to \mathbb{R}$ such that $\mathbb{E}[|h(D_0, \dots, D_n)|^2] < \infty$.

Summarizing the above yields Proposition 4.1.17: A discrete time stochastic process $\{X_n\}$ is a martingale for its canonical filtration if and only if has a zero-mean, orthogonal difference sequence $\{D_n\}$.