Lecture 17: Martingales Examples

Parimal Parag

1 Martingales

Example 1.1. Computing the mean time until a pattern occurs: Consider a sequence of *iid* random variables $\{X_n,\ n \geq 1\}$. Let X_n take values $0, 1$, or 2 with respective probabilities $1/2, 1/3$ and $1/6$. Let N denote the first time the pattern $0, 2, 0$ appears. We are interested in computing $E[N]$. Note that we had solved the same problem earlier using delayed renewal process theory. We shall have a fresh look at the same problem and try to solve it using martingale theory.

Consider a sequence of gamblers , each initially having 1 unit, playing at a fair gambling casino. Gambler i begins betting on i^{th} day and bets the 1 unit he has. If he wins, he bets 2 units on the next outcome being 2. If we wins, he will have 12 units in his hand because he must be earning 10 units for his bet on 2 and the casino is fair $(-2 \times (5/6) + (10) \times (1/6) = 0)$. If he wins the second bet, he bets the 12 units that he earned on the next outcome being 0. At third bet, he quits. At the beginning of each day, a new gambler starts playing. If X_n denotes the total winnings after n^{th} day, then since all bets are fair, it follows that $\{X_n,\ n \geq 1\}$ is a martingale with mean 0. At the end of day N, each of the gamblers $1, 2 \dots N_3$ would have lost 1 unit; gambler $N-2$ would have won 23, gambler $N-1$ would have lost 1, and gambler N would have won 1 (since the outcome on day N is 0). Hence,

$$X_N = N - 3 - 23 + 1 - 1 = N - 26$$

and, since $E[N] = 0$ (Martingale stopping theorem), we see that

$$E[N] = 26.$$

In the same manner, we can find the mean number of coin tosses until a particular pattern, say, $HHTTHH$ occurs to be $p^{-4}q^{-2} + p^{-2} + p^{-1}$, where $p = Pr\{H\} = 1 - q$.

Example 1.2. Coin exchange game: Players X, Y and Z play the following game. At each stage, two of them are chosen randomly in sequence, with the first one chosen being required to give 1 coin to another. All choices are equally and successive choices are independent of the past. This continues until one of the player has no coins and he departs. The game is then continued with the remaining two players until one of them has all the coins. If the players initially has $x,\ y$ and z coins, find the expected number of plays until one of them has all the $s \equiv x + y + z$ coins.

Solution: Let T denote the game ending time. Then T is the first time two players depart. Let $X_n,\ Y_n$ and Z_n denote the amount of money that X, Y and Z have after n^{th} game. Define

$$M_n = X_nY_n + Y_nZ_n + Z_nX_n + n.$$

We will first show that $\{M_n,\ n \geq 1\}$ is a martingale. To that end, observe that M_n is integrable as

$$E[|M_n|] \leq x + y + z + n < \infty.$$

For $X_nY_nZ_n > 0$:

$$\begin{aligned}
E[X_{n+1}Y_{n+1}|X_n, Y_n, Z_n] &= \frac{1}{6}E[(X_n+1)(Y_n-1) + (X_n-1)(Y_n+1) + (X_n)(Y_n+1) + \\
&\quad (X_n)(Y_n-1) + (X_n-1)(Y_n) + (X_n+1)(Y_n)|X_n, Y_n, Z_n] \\
&= X_nY_n - \frac{1}{3}.
\end{aligned}$$

Since other conditional expectations are similarly defined, we find that

$$E[M_{n+1}|X_n, Y_n, Z_n \ldots] = M_n.$$

Now, when $X_nY_nZ_n = 0$, i.e. when one of the players exit the game, say X, then

$$E[Y_{n+1}Z_{n+1}|X_nY_nZ_n] = \frac{1}{2}[(Y_n+1)(X_n-1) + (Y_n-1)(X_n+1)] = Y_nZ_n - 1.$$

Hence, in this case, we again obtain

$$E[M_{n+1}|X_n, Y_n, Z_n, \ldots] = M_n.$$

Hence $\{M_n\}$ is a martingale. By martingale stopping theorem (third condition can be directly verified for the result to hold),

$$E[M_T] = E[M_0].$$

Since two of X_T, Y_T, Z_T are 0,

$$E[M_T] = E[T].$$

Hence,

$$E[T] = E[M_T] = E[M_0] = xy + yz + zx.$$

1.1 Azuma's Inequality for Martingales

Azuma's inequality enables us to obtain useful bounds on the probabilities of a martingale sequence which do not vary too rapidly.

Lemma 1.3. *Let X be a random variable such that $E[X] = 0$ and $Pr(-\alpha \leq X \leq \beta) = 1$. Then for any convex function f*

$$E[f(X)] \leq \frac{\beta}{\alpha+\beta}f(-\alpha) + \frac{\alpha}{\alpha+\beta}f(\beta).$$

Proof. Since $f(x)$ is convex the function value lies below the line segment joining $f(-\alpha)$ and $f(\beta)$. That is,

$$f(x) \leq \frac{\beta}{\alpha+\beta} f(-\alpha) + \frac{\alpha}{\alpha+\beta} f(\beta) + \frac{1}{\alpha+\beta}[f(\beta) - f(-\alpha)]x$$

Taking expectation on both sides gives the required result. □

Lemma 1.4. *For* $0 \leq \theta \leq 1$

$$\theta e^{(1-\theta)x} + (1-\theta)e^{-\theta x} \leq e^{x^2/8}.$$

Proof. Let $\theta = (1+\alpha)/2$ and $x = 2\beta$, we must show that for $-1\alpha \leq +1$,

$$(1+\alpha)e^{\beta(1-\alpha)+(1-\alpha)e^{-\beta(1+\alpha)}} \leq 2e^{\beta^2/2}$$

or, equivalently,

$$e^{\beta} + e^{-\beta} + \alpha(e^{\beta} - e^{-\beta}) \leq 2e^{\alpha\beta+\beta^2/2}.$$

It can be seen that the above inequality is true when $\alpha = -1$ or $\alpha = +1$ and when β is large (say when $|\beta| \leq 100$). Thus, if the lemma were false then the function

$$f(\alpha, \beta) = e^{\beta} + e^{-\beta} + \alpha(e^{\beta} - e^{-\beta}) - 2e^{\alpha\beta+\beta^2/2},$$

would assume a strictly positive maximum in the interior of the region $R = \{(\alpha, \beta) : |\alpha| \leq 1, |\beta| \leq 100\}$. Taking the partial derivatives with respect to α and β and equating to zero, we get,

$$e^{\beta} - e^{-\beta} + \alpha(e^{\beta} + e^{-\beta}) = 2(\alpha+\beta)e^{\alpha\beta+\beta^2/2}$$
$$e^{\beta} - e^{-\beta} = 2\beta e^{\alpha\beta+\beta^2/2}.$$

Assuming a solution in which $\beta \neq 0$ implies, upon division, that

$$1 + \alpha\frac{e^{\beta} + e^{-\beta}}{e^{\beta} - e^{-\beta}} = 1 + \frac{\alpha}{\beta}.$$

By expanding in a power series e^{β} and $e^{-\beta}$, there is no solution for the above equation for $\beta \neq 0,\ \alpha = 0$. Hence, if the lemma is not true, we can conclude that the strictly positive maximal value of $f(\alpha, \beta)$ occurs when $\beta = 0$. However, $f(\alpha, 0) = 0$. Hence the lemma is proven. □

Theorem 1.5. ***Azuma's Inequality*** *Let $Z_n,\ n \geq 1$ be a martingale with mean $\mu = E[Z_n]$. Let $Z_0 = \mu$ and suppose that for nonnegative constants $\alpha_i,\ \beta_i,\ i \geq 1$,*

$$-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i.$$

Then for any $n \geq 0,\ a > 0$:

1. $Pr(Z_n - \mu \geq a) \leq exp\{-2a^2/\sum_{i=1}^{n}(\alpha_i + \beta_i)^2\}$.
2. $Pr(Z_n - \mu \leq -a) \leq exp\{-2a^2/\sum_{i=1}^{n}(\alpha_i + \beta_i)^2\}$.

Suppose first that $\mu = 0$. For $c > 0$

$$Pr(Z_n \geq a) = Pr(exp\{cZ_n\} \geq e^{ca})$$

From Markov inequality, we get

$$Pr(Z_n \geq a) \leq E[exp\{cZ_n\}]e^{-ca}. \tag{1}$$

Let $W_n - expcZ_n$. Note that $W_0 = 0$ and that for $n \geq 1$

$$W_n = exp\{cZ_{n-1}exp\{c(Z_n - Z_{n-1})\}\}.$$

Therefore,

$$\begin{aligned} E[W_n|Z_{n-1}] \quad &= exp\{cZ_{n-1}\}E[exp\{c(Z_n - Z_{n-1})\}|Z_{n-1}] \\ &\leq W_{n-1}[\beta_n exp{-}c\alpha_n + \alpha_n exp\{c\beta_n\}]/(\alpha_n + \beta_n) \end{aligned}$$

where the last inequality follows from Lemma 1.4. Since

1. $f(x) = e^{cx}$ is convex.
2. $-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i$.
3. $E[Z_i - Z_{i-1}|Z_{i-1}] = E[Z_i|Z_{i-1}] - E[Z_i|Z_{i-1}] = 0$.

Taking expectations gives

$$E[W_n] \leq E[W_{n-1}](\beta_n exp\{-c\alpha_n\} + \alpha_n exp\{c\beta_n\})/(\alpha_n + \beta_n).$$

Since $E[W_0] = 1$, iterating the above inequality gives

$$E[W_n] \leq \Pi_{i=1}^{n}\{(\beta_i exp\{-c\alpha_i\}) + \alpha_i exp\{c\beta_i\})/(\alpha_i + \beta_i)\}.$$

Thus from Equation 1, we obtain that for any $c > 0$

$$Pr(Z_n \geq a) \leq e^{-ca}\Pi_{i=1}^{n}\{(\beta_i exp\{-c\alpha_i\} + \alpha_i exp\{c\beta_i\})/(\alpha_i + \beta_i)\} \tag{2}$$

Upon setting $\theta = \alpha_i/(\alpha_i + \beta_i)$ and $x = c((\alpha_i + \beta_i))$ and using Lemma 1.4, we can write

$$Pr(Z_n \geq a) \leq e^{-ca}\Pi_{i=1}^{n} exp\{c^2(\alpha_i + \beta_i)^2/8\}.$$

Thus for any $c > 0$,

$$Pr(Z_n \geq a) \leq exp{-}ca + c^2\sum_{i=1}^{n}(\alpha_i + \beta_i)^2/8.$$

Letting $c = \frac{4a}{\sum_{i=1}^{n}(\alpha_i+\beta_i)^2}$ (which is the value that minimizes the exponent in the above equation) gives that

$$Pr(Z_n \geq a) \leq exp\{-2a^2/\sum_{i=1}^{n}(\alpha_i + \beta_i)^2\}.$$

Parts (i) and (ii) of Azuma's inequality now follow from applying the preceding, first to the zero-mean martingale $\{Z_n - \mu\}$ and to the zero mean martingale $\{\mu - Z_n\}$.

Example 1.6. Suppose that n balls are put in m urns in such a manner that each ball, independently, is equally likely to go into any of the bins. We would like to obtain bounds on the tail probability of the number of empty bins. Let

$$\begin{aligned} X &= \textstyle\sum_{i=1} n1_{\{}\text{bin } i \text{ is empty}\}, \\ E[X] &= mPr(\text{bin } i \text{ is empty}) = m(1-1/m)^n \equiv \mu. \end{aligned}$$

Let X_j denote the bin in which j^{th} ball is placed. Let $Z_0 = E[X]$ and $Z_i = E[X|X_1, X_2 \dots X_i]$. Note that Z_n sequence forms a Doob's martingale sequence. Also, observe that $|Z_1 - Z_0| = 0$. Let D denote the number of distinct values in $X_1 \dots X_{i-1}$. Thus D denotes the number of bins which have at least one ball in it after having distributed $i-1$ balls. Each of the $m-D$ empty bins will end up empty with probability $(1-1/m)^{n-i+1}$. Hence we have

$$E[X|X_1 \dots X_{i-1}] = (m-D)(1-1/m)^{n-i+1}.$$

We have,

$$E[X_i|X_1 \dots X_{i-1}] = \begin{cases} (m-D)(1-1/m)^{n-i}, & \text{if} X_i \in (X_1 \dots X_{i-1}) \\ (m-D-1)(1-1/m)^{n-i}, & \text{if} X_i \notin (X_1 \dots X_{i-1}). \end{cases}$$

Hence, for $i \geq 2$, $Z_i - Z_{i-1}$ are

$$\frac{(m-D)}{m}(1-1/m)^{n-i} \text{ and } \frac{(-D)}{m}(1-1/m)^{n-i}.$$

Since $1 \leq D \leq \min\{i-1, m\}$, we thus obtain that

$$-\alpha_i \leq Z_i - Z_{i-1} \leq \beta_i,$$

where

$$\alpha_i = \min(\frac{i-1}{m}, 1)(1-1/m)^{n-i}, \ \beta_i = (1-1/m)^{n-i+1}.$$

From Azuma's inequality we thus obtain that for $a > 0$

$$Pr(X - \mu \geq a) \leq exp\{-2a^2 / \sum_{i=2}^{n} (\alpha_i + \beta_i)^2\},$$

where

$$\sum_{i=2}^{n} (\alpha_i + \beta_i)^2 = \sum_{i=2}^{n} (m+i-2)^2 (1-1/m)^{2(n-1)}/m^2 + \sum_{i=m+2}^{n} (1-1/m)^2 (2-1/m)^2.$$

Azuma's inequality is generally used in conjunction with Doob's inequality for which $|Z_i - Z_{i-1}| \leq 1$. The following corollory gives a sufficient condition for Doob type martingale to satisfy the condition.

Corollary 1.7. *Let h be a function such that if the vectors $\boldsymbol{x} = (x_1, x_2 \dots x_n)$ and $\boldsymbol{y} = (y_1, y_2 \dots y_n)$ differ in at most one coordinate, then $|h(x) - h(y)| \leq 1$. Let $X_1, X_2 \dots X_n$ be independent random variables. Then, with $\boldsymbol{X} = (X_1 \dots X_n)$ we have for a ¿0 that*

1. *$Pr(h(\boldsymbol{X}) - E[h(\boldsymbol{X})] \geq a) \leq e^{-a^2/2n}$.*

2. $Pr(h(\boldsymbol{X}) - E[h(\boldsymbol{X})] \leq -a) \leq e^{-a^2/2n}$

Proof. Consider the martingale

$$Z_i = E[h(\mathbf{h(X)})|\mathbf{X}_1 \ldots \mathbf{X}_i],\ i = 1, \ldots, n.$$

Now,

$$\begin{aligned} &|E[h(\mathbf{h(X)})|\mathbf{X}_1 \ldots \mathbf{X}_i] - E[h(\mathbf{h(X)})|\mathbf{X}_1 \ldots \mathbf{X}_{i-1}]| \\ &= |E[h(x_1, \ldots x_i, X_{i+1}, \ldots, X_n)] - E[h(x_1, \ldots x_{i-1}, X_i, \ldots, X_n)]| \\ &= |E[h(x_1, \ldots x_i, X_{i+1}, \ldots, X_n) - h(x_1, \ldots x_{i-1}, X_i, \ldots, X_n)]| \end{aligned}$$

Hence, $|Z_i - Z_{i-1}| \leq 1$ and so the result follows from Azuma's inequality with $\alpha_i = \beta_i = 1$. $\square$