5.4. **Optimal Sampling Theorem (OST).** First I stated it a little vaguely:

Theorem 5.12. *Suppose that*

(1) *T is a stopping time*
(2) *M_n is a martingale wrt the filtration $\mathcal{F}_n$*
(3) *certain other conditions are satisfied.*

Then:

$$\boxed{\mathbb{E}(M_T \,|\, \mathcal{F}_0) = M_0}$$

The first thing I explained is that this statement is NOT TRUE for Monte Carlo. This is the gambling strategy in which you double your bet every time you lose. Suppose that you want to win \$100. Then you go to a casino and you bet \$100. If you lose you bet \$200. If you lose again, you bet \$400 and so on. At the end you get \$100. The probability is zero that you lose every single time. In practice this does not work since you need an unlimited supply of money. But in mathematics we don't have that problem.

To make this a martingale you do the following. Let

$$X_1, X_2, X_3, \cdots$$

be i.i.d. Bernoulli random variables which are equal to ±1 with equal probability:

$$X_i = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases}$$

In other words, we are assuming each gave is fair. Then

$$\mathbb{E}(X_i) = 0.$$

Let

$$M_n = X_1 + 2X_2 + 4X_3 + \cdots + 2^{n-1}X_n$$

This is the amount of money you will have at the end of n rounds of play if you bet 1 on the first game, 2 on the second, 4 on the third, etc. and keep playing regardless of whether you win or lose. To see that this is a martingale we calculate:

$$M_{n+1} = X_1 + 2X_2 + \cdots + 2^{n-1}X_n + 2^n X_{n+1} = M_n + 2^n X_{n+1}$$

At time n we know the first n numbers but we don't know the last number. So,

$$\begin{aligned} \mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n) &= M_n + \mathbb{E}(2^n X_{n+1}) \\ &= M_n + 2^n \mathbb{E}(X_{n+1}) = M_n + 0 = M_n \end{aligned}$$

I.e., the expect future value is the same as the known value on each day. So, this is a martingale.

$T =$ the first time you win. Then

$$\mathbb{P}(T < \infty) = 1.$$

The argument about random walk being null recurrent actually does not apply here. I will explain on Monday what that was about. In the Monte Carlo case it is obvious that $T < \infty$ since

$$\mathbb{P}(T > n) = \frac{1}{2^n} \to 0.$$

In any case,

$$M_T = 1$$

since, at the moment you win, your net gain will be exactly 1. So,

$$\mathbb{E}(M_T \,|\, \mathcal{F}_0) = 1 \neq M_0 = 0.$$

In other words, the Optimal Sampling Theorem does not hold. We need to add a condition that *excludes Monte Carlo.* We also know that we cannot prove a theorem which is false. So, we need some other condition in order to prove OST. The simplest condition is boundedness:

Theorem 5.13 (OST1)**.** *The OST holds if T is bounded, i.e., if $T \leq B$ for some constant B.*

5.5. **integrability conditions.** The OST says that

$$\mathbb{E}(M_T \mid \mathcal{F}_0) = M_0$$

under "certain conditions." These are integrability conditions which I want to explain (but just the definition).

Definition 5.14. Suppose that Y is a random variable. Then

(1) Y is *integrable* (L^1) if

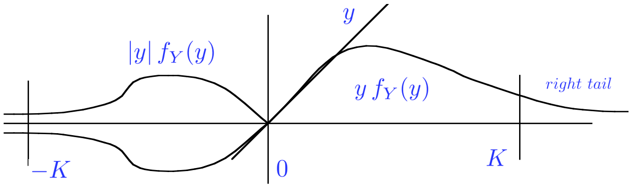
$$\mathbb{E}(|Y|) < \infty$$

I.e, if the integral

$$\int_{-\infty}^{\infty} |y| f_Y(y) dy \quad \text{converges.}$$

(2) Y is *square integrable* (L^2, p. 217 in book) if

$$\mathbb{E}(Y^2) < \infty$$



If Y is *not* integrable then one of the tails must be fat: I.e., either the right tail

$$\int_K^{\infty} y\, f_Y(y) dy = \infty$$

or the left tail:

$$\int_{-\infty}^{-K} |y|\, f_Y(y) dy = \infty$$

If we cut off any finite piece, you still have infinity left. So, the same will be true for any value of K.

5.5.1. *a beautiful theorem.* Here is a wonderful theorem which takes longer to state than to prove and which is related to what we just learned.

Theorem 5.15. *Suppose that*

(1) Y_n *is* $\mathcal{F}_n$*-measurable,*
(2) T *is a stopping time and*
(3) $\mathbb{P}(T < \infty) = 1$.

Then

$$M_n := \mathbb{E}(Y_T \,|\, \mathcal{F}_n)$$

is a martingale wrt $\mathcal{F}_n$.

Proof.

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(Y_T \,|\, \mathcal{F}_{n+1}) \,|\, \mathcal{F}_n) = \mathbb{E}(Y_T \,|\, \mathcal{F}_n) = M_n.$$

So, M_n is a martingale. □

Example 5.16. Let $Y_n = f(X_n)$ be the payoff function.
$X_n =$ state at time n.
$T =$ optimal stopping time. Then
$Y_T = f(X_T) =$ optimal payoff.
$v(x) =$ value function.
Then

$$v(X_n) = \mathbb{E}(\underbrace{f(X_T)}_{Y_T} \,|\, \mathcal{F}_n)$$

As an example of the theorem we just proved, we have:

Corollary 5.17. $M_n = v(X_n)$ *is a martingale!*

Question: Does $v(X_n)$ satisfy OST? In other words:

$$¿\mathbb{E}(v(X_T) \,|\, \mathcal{F}_0) = v(X_0)?$$

Answer: Yes, because $v(X_T) = f(X_T)$. (When you reach the state X_T you are supposed to stop and take the payoff.)

5.5.2. *uniform integrability.*

Theorem 5.18 (2nd Optimal Sampling Theorem)**.** *Suppose that* $M_0, M_1, M_2, \cdots$ *is a martingale wrt the filtration* $\mathcal{F}_n$. *Suppose*

1. $T =$ *stopping time*
2. $\mathbb{P}(T < \infty) = 1$.
3. M_T *is integrable* $\mathbb{E}(|M_T|) < \infty$
4. $M_0, M_1, \cdots$ *are uniformly integrable (defined below).*

Then OST holds, i.e., $\mathbb{E}(M_T \,|\, \mathcal{F}_0) = M_0$.

Note: The contrapositive is also true. I.e., if OST fails then one of the conditions must fail.

For example, in Monte Carlo, $X_i = \pm 1$ with probability $1/2$,

$$M_n = X_1 + 2X_2 + 2^2 X_3 + \cdots + 2^{n-1} X_n$$

is a martingale

$$T = \text{ smallest } n \text{ so that } X_n = 1.$$

This is a stopping time with $\mathbb{P}(T < \infty) = 1$ and $M_T = 1$ is integrable. But OST fails. So, it must be that this martingale is not uniformly integrable.

Definition 5.19. Y_n is integrable if for every $\epsilon > 0$ there is a $K_n > 0$ so that the K_n-tails have total area less than ϵ:

$$\int_{K_n}^{\infty} y f_{Y_n}(y)dy + \int_{-\infty}^{-K_n} |y| f_{Y_n}(y)dy < \epsilon$$

Y_n is *uniformly integrable* if the cutoff points are the same for all Y_n: $K_n = K$.

If a sequence Y_n is not uniformly integrable then, as time goes on, you are very likely to end up in the tail. (No matter where you cut it the tail has probability $\geq \epsilon > 0$. But you have an infinite sequence of random variable. If they are independent you are almost certain to end up in the tail.)

Finally, I asked: Why is Monte Carlo not uniformly integrable? It is not given by an integral. So, what does this mean?

5.5.3. *nonintegral meaning of uniform integrability.* We need a new definition of tail which applies to any random variable Y_n, not just the continuous ones. For any $\delta > 0$ define a *δ-tail* to be a set of values of Y_n with probability $\leq \delta$. Then uniform integrability implies that: $\forall \epsilon > 0\ \exists \delta > 0$ so that

$$\int_{\delta\text{-tail}} |Y_n| < \epsilon$$

for all n. (In the discrete case the integral means you add up the probability times $|Y_n|$ for all points in the tail.) In the case of Monte Carlo, regardless of δ, we can take n so that $1/2^n < \delta$. Then the event that $X_1, X_2, \cdots, X_n$ are all 1 is in the δ-tail. It has probability $1/2^n$. But $M_n = 2^n - 1$ on this tail. So,

$$\int_{\delta\text{-tail}} |M_n| \geq \frac{2^n - 1}{2^n} \approx 1$$

which will not be $< \epsilon$. So, this sequence is not uniformly integrable.

This δ-tail condition is not exactly the same uniform integrability. This will be explained at the end.

5.5.4. *Martingale convergence theorem.* I just stated this theorem without much explanation. It has two important integrality conditions.

Theorem 5.20 (Martingale convergence theorem). *Suppose that M_n is a martingale wrt the filtration $\mathcal{F}_n$. Then*

(1) *M_n converges to a random variable M_∞ if $\mathbb{E}(|M_n|) \leq C$ for some constant C.*
(2) *$\mathbb{E}(M_n) \to \mathbb{E}(M_\infty)$ if M_n are uniformly integrable.*

This ends what I said in class about martingales. What follows are some theoretical comments that I didn't have time to say. If we need them later I will go back and explain them.

It helps to know that the second conditions implies the first condition.

Lemma 5.21. *If Y_n are uniformly integrable then there is finite C so that $\mathbb{E}(|Y_n|) \leq C$ for all n.*

In fact there is the following theorem relating uniform integrability, this boundedness condition and the δ-tail interpretation.

Theorem 5.22. *A sequence of real valued random variables Y_n is uniformly integrable if and only if both of the following conditions hold.*

(1) *(uniform L^1-boundedness) $\exists C < \infty$ s.t. $\mathbb{E}(|Y_n|) \leq C$ for all n*
(2) *(δ-tail condition) $(\forall \epsilon > 0)(\exists \delta > 0)$*

$$\int_{\delta\text{-}tail} |Y_n| < \epsilon$$

for all n.

5.5.5. *definition of uniform integrability.* The book gives the following definition of uniform integrability. This wording is intended to apply to all cases of real valued random variables.

Definition 5.23. A sequence or real valued random variables Y_n is *uniformly integrable* iff $(\forall \epsilon > 0)(\exists K > 0)$ so that

$$\mathbb{E}(|Y_n| I(|Y_n| > K) \leq \epsilon$$

Where $I(|Y_n| > K)$ is the indicator function of the property $|Y_n| < K$, i.e., it is the function which is equal to 1 when $|Y_n| < K$ and 0 elsewhere. Expectation value are given by integrals for continuous random variables and sum for discrete random variables. So, this is always defined.

Proof of Lemma 5.21. Another one-line proof:

$$\mathbb{E}(|Y_n|) = \mathbb{E}(|Y_n| I(|Y_n| \leq K) + \mathbb{E}(|Y_n| I(Y > K)) \leq 2K^2 + \epsilon.$$

□