October 24 495 Lecture

I. Review

At the end of the previous lecture we stated and applied the formula

$$E[X] = E\big[E[X|Y]\big] \qquad (1)$$

for two discrete random variables. To recall, for each value y such that $\mathbb{P}(Y=y)$

- $$E[X|Y=y] \overset{\text{def}}{=} \sum_x x\,\mathbb{P}(X=x|Y=y) \qquad (2)$$
- $E[X|Y]$ is the random variable obtained whose value when $Y=y$ is $E[X|Y=y]$

Then, by definition of expected value, the right hand side of (1) is

$$E\big[E[X|Y]\big] = \sum_y E[X|Y=y]\,\mathbb{P}(Y=y). \qquad (3)$$

The left hand side of (1) is

$$E[X] = \sum_x x\,\mathbb{P}(X=x)$$

To see why (1) is true, plug (2) into (3) and recall that $\mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) = \mathbb{P}(X=x, Y=y)$ and

$$\mathbb{P}(X=x) = \sum_y \mathbb{P}(X=x, Y=y)$$

$$\begin{aligned} E\big[E[X|Y]\big] &= \sum_y \Big(\sum_x x\,\mathbb{P}(X=x|Y=y)\Big)\mathbb{P}(Y=y) \\ &= \sum_y \sum_x x\,\mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) \\ &= \sum_x x \sum_y \mathbb{P}(X=x, Y=y) \\ &= \sum_x x\,\mathbb{P}(X=x) = E[X] \end{aligned}$$

II. IMPORTANT GENERALIZATIONS OF (1)

Let X and $Y_1, \dots, Y_n$ be discrete random variables. Define,

$$\left.\begin{aligned} &E[X|Y_1=y_1,\dots,Y_n=y_n] \overset{\text{def}}{=} \sum_x x\,\mathbb{P}(X=x|Y_1=y_1,\dots,Y_n=y_n) \\ &\qquad\text{for all } y_1 \dots y_n \text{ such that } \mathbb{P}(Y_1=y_1,\dots,Y_n=y_n)>0 \end{aligned}\right\} (4)$$

This definition generalizes (1)

Define

$$E[X|Y_1,\dots,Y_n]$$

to be the random variable which takes the value (4) when the random variable $Y_1=y_1$, the r.v. $Y_2=y_2$, …, the r.v. $Y_n=y_n$.

October 24 **495 Lecture**

I Review

At the end of the previous lecture we stated and applied the formula

$$E[X] = E\big[E[X|Y]\big] \qquad (1)$$

for two discrete random variables. To recall, for each value y such that $\mathbb{P}(Y=y)$

- $$E[X|Y=y] \overset{\text{def}}{=} \sum_x x\,\mathbb{P}(X=x \mid Y=y) \qquad (2)$$
- $E[X|Y]$ is the random variable obtained whose value when $Y=y$ is $E[X|Y=y]$

Then, by definition of expected value, the right hand side of (1) is

$$E\big[E[X|Y]\big] = \sum_y E[X|Y=y]\,\mathbb{P}(Y=y). \qquad (3)$$

The left hand side of (1) is

$$E[X] = \sum_x x\,\mathbb{P}(X=x)$$

To see why (1) is true, plug (2) into (3) and recall that $\mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) = \mathbb{P}(X=x, Y=y)$ and

$$P(X=x) = \sum_y \mathbb{P}(X=x, Y=y)$$

$$\begin{aligned}
E\big[E[X|Y]\big] &= \sum_y \Big(\sum_x x\,\mathbb{P}(X=x|Y=y)\Big)\mathbb{P}(Y=y) \\
&= \sum_y \sum_x x\ \mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) \\
&= \sum_x x \sum_y \mathbb{P}(X=x, Y=y) \\
&= \sum_x x\,\mathbb{P}(X=x) = E[X]
\end{aligned}$$

II IMPORTANT GENERALIZATIONS OF (1)

Let X and $Y_1, \dots, Y_n$ be discrete random variables. Define,

$$\left.\begin{aligned} E[X|Y_1=y_1,\dots,Y_n=y_n] &\overset{\text{def}}{=} \sum_x x\,\mathbb{P}(X=x \mid Y_1=y_1,\dots,Y_n=y_n) \\ &\text{for all } y_1,\dots,y_n \text{ such that } \mathbb{P}(Y_1=y_1,\dots,Y_n=y_n) > 0 \end{aligned}\right\} (4)$$

This definition generalizes (1)

Define

$$E[X|Y_1,\dots,Y_n]$$

to be the random variable which takes the value (4) when the random variable $Y_1=y_1$, the r.v. $Y_2=y_2$, .., the r.v. $Y_n=y_n$.

By definition of expected value

$$E\left[E[X|Y_1,\dots,Y_n]\right] = \sum_{y_1,\dots,y_n} E[X|Y_1=y_1,\dots,Y_n=y_n]\,\mathbb{P}(Y_1=y_1,\dots,Y_n=y_n). \quad (5)$$

(The sum is over all possible values $y_1 \cdots y_n$ of $Y_1 \cdots Y_n$.)

The following identity is true and generalizes (1)

$$\boxed{E[X] = E\left[E[X|Y_1,\dots,Y_n]\right] \qquad (6)}$$

(6) is demonstrated in the same way we demonstrated (1). Only the notation is more complicated. We need only observe that

$$\mathbb{P}(X=x) = \sum_{y_1,\dots,y_n} \mathbb{P}(X=x, Y_1=y_1,\dots,Y_n=y_n)$$

$$= \sum_{y_1,\dots,y_n} \mathbb{P}(X=x|Y_1=y_1,\dots,Y_n=y_n)\,\mathbb{P}(Y_1=y_1,\dots,Y_n=y_n)$$

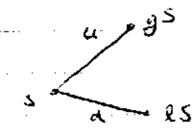
So if (4) is substituted in (5),

$$E\left[E[X|Y_1,\dots,Y_n]\right] = \sum_{y_1\cdots y_n}\left(\sum_x x\,\mathbb{P}(X=x|Y_1=y_1,\dots,Y_n=y_n)\right)\mathbb{P}(Y_1=y_1,\dots,Y_n=y_n)$$

$$= \sum_x x \sum_{y_1\cdots y_n} \mathbb{P}(X=x|Y_1=y_1,\dots,Y_n=y_n)\,\mathbb{P}(Y_1=y_1,\dots,Y_n=y_n)$$

$$= \sum_x x\,\mathbb{P}(X=x) = E[X]$$

Example Consider the binomial tree model with the risk-neutral measure. Let g be the factor by which the price increases in one period if u occurs; let ℓ be the factor by which it decreases if d occurs and let $\tilde{p} = (e^{r\tau}-\ell)/g-\ell$, $\tilde{q} = 1-\tilde{p}$ be the risk neutral probabilities of u and ℓ. Recall that for the risk neutral measure market movements in different periods are independent.



This model says

$$S_{t+1} = \begin{cases} g S_t & \text{if } u \text{ occurs in period } t+1 \\ \ell S_t & \text{if } d \text{ occurs in period } t+1. \end{cases}$$

What is $\tilde{E}[S_{t+1}|S_1,\dots,S_t]$?

Apply definition (4)

$$\tilde{E}[S_{t+1}|S_1=s_1,\dots,S_t=s_t]$$

$$= \sum_s s\,\mathbb{P}(S_{t+1}=s|S_1=s_1,\dots,S_t=s_t)$$

By the model, since whether the market goes up or down in period $t+1$ is independent of $S_1 \cdots S_t$.

$$\mathbb{P}[S_{t+1}=\ell s_t|S_1=s_1,\dots,S_t=s_t] = \tilde{q}$$

$$\mathbb{P}[S_{t+1}=g s_t|S_1=s_1,\dots,S_t=s_t] = \tilde{p}$$

and, conditioned on $S_1=s_1,\dots,S_t=s_t$, the only possible values of S_{t+1} are ℓs_t and $g s_t$. Thus

$$\tilde{E}[S_{t+1} \mid S_1 = s_1, \ldots, S_t = s_t] = \ell s_t \tilde{q} + g s_t \tilde{p}$$

$$= s_t(\ell\tilde{q} + g\tilde{p}) = s_t e^{r\tau} \quad (7)$$

since $\tilde{q}$ and $\tilde{p}$ are defined by the condition $\ell\tilde{q} + g\tilde{p} = e^{r\tau}$.

Notice in (7) that the conditional expectation of S_{t+1} given $S_1 = s_1, \ldots, S_t = s_t$ depends only on the last price $S_t = s_t$ -- the other prices are irrelevant.

(7) implies

$$\tilde{E}[S_{t+1} / S_1, \ldots S_t] = S_t e^{r\tau} \quad (8)$$

since $\tilde{E}[S_{t+1} \mid S_1, \ldots, S_t]$ is obtained from (7) by replacing $s_1, \ldots, s_t$ by the random values $S_1, \ldots, S_t$.

Let us now apply (6)

$$\tilde{E}[S_t] = \tilde{E}[\; E[S_t \mid S_1, \ldots, S_{t-1}]\;]$$

$$= \tilde{E}[S_{t-1} e^{r\tau}] = e^{r\tau}\tilde{E}[S_{t-1}] \quad (9)$$

Iterated (9). Then

$$\tilde{E}[S_t] = e^{r\tau}\tilde{E}[S_{t-1}] = e^{r\tau} \cdot e^{r\tau}\tilde{E}[S_{t-2}]$$

$$= \ldots$$

$$= \underbrace{e^{r\tau} \cdot e^{r\tau} \cdots e^{r\tau}}_{t \text{ times}} \tilde{E}[S_0]$$

$$= (e^{r\tau})^t S_0$$ (since S_0 is given and not random)

IMPORTANT REMARK

If $S_1, \ldots, S_t$ is known then the market history up to time t is known because we can read off whether an upswing or downswing occurred from looking at how S_t moves. For example if $S_1 = \ell S_0$, $S_2 = \ell^2 S_0$, $S_3 = g\ell^2 S_0$, $S_4 = g^2\ell^2 S_0$ we know $\omega_1 = d$, $\omega_2 = d$, $\omega_3 = u$, $\omega_4 = u$.

In previous lectures, we used $\tilde{E}[V \mid \mathcal{F}_t]$ to denote the conditional expectation ~~when~~ of V when we conditioned on the whole market history up to time t. Since the information in $S_1, \ldots, S_t$ is the same as the market history up to time t, we have

$$\boxed{\tilde{E}[V \mid \mathcal{F}_t] = \tilde{E}[V \mid S_1, \ldots, S_t] \quad (10)}$$

as an identity of random variables (Even if you don't see this by formal proof going back to definitions, you should understand (and accept) (10) and be able to use it).

By applying (10) in (6) we obtain

$$\boxed{\tilde{E}[V] = \tilde{E}\left[\tilde{E}[V|\mathcal{F}_t]\right] \qquad (11)}$$

for all t. This is the formula we stated last time and used to show that

if V_0 is computed from V by backward induction algorithm, meaning

$$(12)\quad \begin{cases} V_N(w_1,\dots w_N) = V(w_1 \dots w_N), \text{ all } w_1 \dots w_N \\ V_t(\omega_1,\dots,w_t) = \frac{1}{e^{r\tau}}\left[\tilde{q}V_{t+1}(\omega_1,\dots,w_t,d) + \tilde{p}V_{t+1}(w_1,\dots,w_t,u)\right] \\ t = N-1, N-2, \dots, 0 \end{cases}$$

then

$$\boxed{V_0 = (e^{-r\tau})^N \tilde{E}[V_N] \qquad (13)}$$

For the record, we repeat the derivation of (13): it is like (exactly like) the derivation of $E[S_N] = (e^{r\tau})^N S_0$. The crucial point is to observe that the induction

step in (12) is really the same as

$$V_t(\omega_1,\dots,\omega_t) = e^{-r\tau}\tilde{E}[V_{t+1}|\omega_1 \dots \omega_t];$$

in other words

$$(14)\qquad V_t = e^{-r\tau}\tilde{E}[V_{t+1}|\mathcal{F}_t] \qquad t = 0,1,\dots,N-1$$

Hence

$$\begin{aligned} \tilde{E}[V_N] &= \tilde{E}\left[\tilde{E}[V_N|\mathcal{F}_{N-1}]\right] \\ &= e^{r\tau}\tilde{E}[V_{N-1}] \qquad \text{using (14) with } t = N-1 \\ &= e^{r\tau}e^{r\tau}\tilde{E}[V_{N-2}] \\ &= \dots\dots \\ &= \underbrace{e^{r\tau}\dots e^{r\tau}}_{N \text{ times}}\tilde{E}V_0 \\ &= (e^{r\tau})^N V_0 \end{aligned}$$

Thus $\qquad V_0 = (e^{-r\tau})^N \tilde{E}[V_N]$

This ends this long example. You are asked to explore (5) and (6) further in problems 49-52.

∎

Next we come to a second important generalization of (1) and (6). This is the following:

If $1 \le k < n$ then

$$E[X \mid Y_1, \dots, Y_k] = E\left[E[X \mid Y_1 \dots Y_k, Y_{k+1} \dots Y_n] \mid Y_1 \dots Y_k \right] \quad (15)$$

This generalizes (6) in the sense that the expectation in (6) is replaced by conditional expectation on $Y_1, \dots Y_k$. It is very important to have $n > k$ in (15); for example

$$E[X \mid Y_1, Y_2] \neq E\left[E[X \mid Y_1] \mid Y_1, Y_2 \right]$$

in general

The reason for (15) is the same as for (6). To prove (15) one follows the steps proving (6) but conditioning on $Y_1, \dots Y_k$ throughout. You are asked to do a special case in problem 53.

Example: Applying (15) to the binomial tree with risk neutral probabilities.

We continue in the notation of the previous example.

(a) Using (15)

$$\tilde{E}[S_4 \mid S_1, S_2] = \tilde{E}\left[\tilde{E}[S_4 \mid S_1, S_2, S_3] \mid S_1, S_2 \right]$$

$$= \tilde{E}[e^{r\tau} S_3 \mid S_1, S_2] \quad \text{(using (8))}$$

$$= e^{r\tau} e^{r\tau} S_2 \quad \text{(using (8) again)}.$$

$$= (e^{r\tau})^2 S_2. \quad (16)$$

This is a special example of the formula

$$\tilde{E}[S_t \mid S_1, \dots S_j] = (e^{r\tau})^{t-j} S_j \quad (17)$$

which the student can show by imitating the derivation of (16)

(b) Very important application!.

Let $V_t(\omega_1, \dots, \omega_t)$ be obtained from V by the backward induction formula in (12). We have been using $V_t(\omega_1, \dots \omega_t)$ as the claimed no arbitrage price ~~of the deriv~~ at time t if $\omega_1 \dots \omega_t$ has occured of a derivative paying V at time N.

Claim

$$\boxed{V_t = (e^{r\tau})^{N-t} \tilde{E}[V_N \mid \mathcal{F}_t] \quad (18)}$$

NO ARBITRAGE PRICES ARE OBTAINED AS CONDITIONAL EXPECTATIONS WITH RESPECT TO THE RISK NEUTRAL PROBABILITY MEASURE.

Demonstration. Since $\tilde{E}[V_N | \mathcal{F}_t] = \tilde{E}[V | S_1, \ldots S_t]$ we show, ~~that~~

$$V_t = (e^{-r\tau})^{N-t} \tilde{E}[V_N | S_1, \ldots S_t]$$

Again, we use that the induction equation is

$$V_t = e^{-r\tau} \tilde{E}[V_{t+1} | S_1, \ldots, S_t] \quad (19)$$

Thus

$$\tilde{E}[V_N | S_1, \ldots S_t] = \tilde{E}\Big[\tilde{E}[V_N | S_1, \ldots, S_{N-1}] \Big| S_1, \ldots S_t\Big] \quad (\text{by } (15))$$

$$= \tilde{E}[e^{r\tau} V_{N-1} | S_1, \ldots, S_t] \quad (\text{by } (19))$$

$$= e^{r\tau} \tilde{E}\Big[E[V_{N-1} | S_1 \ldots S_{N-2}] \Big| S_1 \ldots S_t\Big] \quad (\text{by } (15))$$

$$= (e^{r\tau})^2 \tilde{E}[V_{N-2} | S_1, \ldots S_t] \quad \text{by } (19).$$

Continue this procedure, we get

$$\tilde{E}[V_N | S_1 \ldots S_t] = (e^{r\tau})^{N-t+1} \tilde{E}[V_{t+1} | S_1 \ldots S_t]$$

$$= (e^{r\tau})^{N-t} V_t \quad (\text{by } (19))$$

Dividing both sides by $(e^{r\tau})^{N-t}$ gives (18).

III. Martingales

The conditional expectation machinery allows us to define martingales.

Let $Y_0, Y_1, Y_2, Y_3, \ldots$ be a sequence of random variables. We think of $Y_0, Y_1, \ldots, Y_t$ as the information available to an observer up to time t. For example in the binomial tree model $Y_1, \ldots Y_t$ might be the prices $S_1, \ldots, S_t$ which encode the market history up to time t.

Let $X_0, X_1, X_2, \ldots$ be a second sequence of random variables. We are definitely thinking of the index k in X_k as a time index here, so think of $X_1, X_2, \ldots$ as random variables being observed successively in time.

Definition $X_0, X_1, X_2, \ldots$ is a martingale w.r.t. the sequence $\mathcal{Y} = \{Y_0, Y_1, Y_2, Y_3, \ldots\}$ if

(i) $E|X_n| < \infty$ for each n

(ii) For each n X_n is determined by $Y_1, \ldots, Y_n$ -- that is, if $Y_1 \cdots Y_n$ are known then X_n is known so X_n is effectively a function of $Y_1 \cdots Y_n$

(iii) $$E[X_{n+1} | Y_1, \ldots Y_n] = X_n \quad \text{for every } n.$$

Condition (iii) is the minimum requirement one would make for $\{X_n\}$ to be a model of a fair game.

Examples a) Let $Y_0, Y_1, Y_2, \dots$ be independent random variables such that for each i

$$\mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i = -1) = 1/2$$

(Think of Y_i as an amount won or lost on play i)

Note that $E[Y_i] = 0$ for each i and

$$E[Y_{i+1} | Y_1 \dots Y_i] = E[Y_i] \quad \text{(by independence)}$$
$$= 0 \qquad (20)$$

for each i.

Let x be a real number and define $X_0 = x$, $X_1 = x + Y_1$, $X_2 = x + Y_1 + Y_2, \dots, X_n = x + \sum_1^n Y_i, \dots$

$\{X_n\}_{n \geq 1}$ is the classical, symmetric random walk starting from x. We claim that $\{X_n\}$ is a martingale relative to $\{Y_n\}$. To see this, we check conditions (i), (ii), (iii) of the definition. Condition (i) is easy because $-n + x \leq X_n \leq x + n$ for each n; hence $|X_n|$ is bounded and so $E|X_n|$ must be finite. Since $X_n = x + \sum_1^n Y_i$ (ii) is satisfied. For (iii)

$$E[X_{n+1} | Y_1 \cdots Y_n] = E[x + \sum_1^n Y_i + Y_{n+1} | Y_1 \cdots Y_n]$$

$$= x + \sum_1^n Y_i + E[Y_{n+1} | Y_1 \cdots Y_n]$$

$$= x + \sum_1^n Y_i + 0 \quad (\text{by } (20))$$

$$= X_n .$$

(Note: Y_0 plays no role in $X_0, X_1, \dots$ It is kept here for consistency with the definition of martingale above.)

b) Geometric random walk

Let $Y_0, Y_1, Y_2, \dots$ be independent and assume

$$\mathbb{P}(Y_i = 1) = p \qquad \mathbb{P}(Y_i = -1) = q$$

$Z_n = \sum_1^n Y_i$, $n \geq 1$, again defines a random walk, (although if $p \neq 1/2$ it is no longer a martingale).

Let $\lambda > 0$ and $\mu > 0$ be constants. A process of the form $X_0 = A$

$$X_n = A \mu^n \lambda^{Y_1 + \cdots + Y_n} \qquad n \geq 1$$

is called a geometric random walk.

We claim that if

$$\mu = \frac{1}{\lambda^{-1} q + \lambda p}$$

then $\{X_n\}$ is a martingale w.r.t. $\{Y_n\}$.

Again (i) and (ii) in the martingale definition are easy to verify. As for (iii), note

$$E[X_{n+1} | Y_1, \dots, Y_n] = E[A \mu^{n+1} \lambda^{Y_1 + \cdots + Y_n} \lambda^{Y_{n+1}} | Y_1 \cdots Y_n]$$

$$= A \mu^{n+1} \lambda^{Y_1 + \cdots + Y_n} E[\lambda^{Y_{n+1}} | Y_1 \cdots Y_n]$$

(because once $Y_1, \dots, Y_n$ $A \mu^{n+1} \lambda^{Y_1 + \cdots + Y_n}$ is a known constant)

$$= A \mu^{n+1} \lambda^{Y_1 + \cdots + Y_n} E[\lambda^{Y_{n+1}}]$$

(because Y_n is independent of $Y_1, \dots, Y_n$)

$$= X_n \mu \left(q\lambda^{-1} + p\lambda \right) \qquad \text{since } E[\lambda^{Y_{n+1}}] = q\lambda^{-1} + p\lambda$$

$$= X_n \qquad \text{since } \mu = \frac{1}{q\lambda^{-1} + p\lambda}$$

c) Under the risk-neutral measure on binomial trees $\{S_t\}$ is a geometric random walk and the discounted price process

$$(e^{-r\tau})^t S_t$$

is a geometric random walk and a martingale.

Let $Y_i = 1$ if u occurs in period i and $Y_i = -1$ if d occurs in i.

In t periods

$$\sum_1^t Y_i = (\#\text{ of times } u \text{ occurs}) - (\#\text{ of times } d \text{ occurs})$$

$$= (\#\text{ of times } u \text{ occurs}) - (t - (\#\text{ of times } u \text{ occurs}))$$

$$= 2(\#\text{ of times } u \text{ occurs}) - t$$

Hence

$$\#\text{ of times } u \text{ occurs up to } t = \frac{\sum_1^t Y_i + t}{2} \qquad (21)$$

Similarly

$$\#\text{ of times } d \text{ occurs up to } t = \frac{t - \sum_1^t Y_i}{2}$$

Thus

$$S_t = S_0\, g^{\frac{\sum_1^t Y_i + t}{2}}\, \ell^{\frac{t - \sum_1^t Y_i}{2}} = S_0 (\sqrt{g\ell})^t \left(\sqrt{\frac{g}{\ell}}\right)^{\sum_1^t Y_i} \qquad (22)$$

Under the risk-neutral measure $Y_1, Y_2, \ldots$ are independent and $\mathbb{P}(Y_i = 1) = \tilde{p}$, $\mathbb{P}(Y_i = -1) = \tilde{q}$ so it follows from (22) that $\{S_t\}$ is a geometric random walk.

We could use the calculation of (6) to show $(e^{-r\tau})^t S_t$ is a martingale under the risk-neutral measure. But we don't need to because we have essentially already shown this martingale property. Indeed, we derived (8) which says:

$$\tilde{E}\left[S_{t+1} \middle| S_1, \ldots S_t \right] = e^{r\tau} S_t$$

Multiply both sides by $(e^{-r\tau})^{t+1}$ to derive

$$\tilde{E}\left[(e^{-r\tau})^{t+1} S_{t+1} \middle| S_1 \cdots S_t \right] = (e^{-r\tau})^t S_t. \qquad (23)$$

This is the martingale property we are claiming (Really we should write $\tilde{E}[\,\cdot \mid Y_1, \ldots Y_t]$ in this formula instead of $\tilde{E}[\,\cdot \mid S_1 \ldots S_t]$ but $S_1, \ldots, S_t$ and $Y_1 \cdots Y_t$ contain the same information and so $\tilde{E}[X \mid S_1 \cdots S_t] = \tilde{E}[X \mid S_1 \cdots S_t]$ for any r.v. X.)

<u>Claim</u> (that we will not prove)

The only probability measure on the market history paths of the binomial tree model that makes $(e^{-r\tau})^t S_t$ a martingale is the risk-neutral measure.

d) VERY IMPORTANT

Consider a derivative with payoff V_N at time N. Let $V_0, V_1, \ldots, V_{N-1}$ be the prices computed by the backward induction algorithm (12). Then $\{(e^{-r\tau})^t V_t\}$ is a martingale w.r.t. $\{S_t\}$ under the risk-neutral measure (Because of (10), which says $\tilde{E}[V|\mathcal{F}_t] = \tilde{E}[V|S_1 \cdots S_t]$, we also say $\{(e^{-r\tau})^t V_t\}$ is a martingale relative to $\{\mathcal{F}_t\}$.)

This fact is by definition! Recall that (12) is the same as (14), which states

$$V_t = e^{-r\tau}\tilde{E}[V_{t+1}|\mathcal{F}_t] = e^{-r\tau}\tilde{E}[V_{t+1}|S_1, \ldots, S_t]$$

Multiply both sides by $(e^{-r\tau})^t$. Then

$$(e^{-r\tau})^t V_t = \tilde{E}[(e^{-r\tau})^{t+1} V_{t+1} | S_1 \cdots S_t] \quad (24)$$

This is precisely the martingale property of $\{(e^{-r\tau})^t V_t\}$.