6.262 LECTURE 8 3/2/09

- Stopping times and Wald's equality

- Elementary renewal theorem

- Blackwell's theorem

- Residual life

Random stopping times

Let J be a positive integer rv and consider observing $X_1, \ldots, X_J$, the first J rv's out of an IID sequence $\{X_i; i \geq 1\}$.

At time 1, $X_1(\omega)$ is observed.

At time 2, a decision, based on $X_1(\omega)$, is made to either observe $X_2(\omega)$ or to stop with $J(\omega) = 1$.

At time 3, if $X_2(\omega)$ was observed, a decision, based on $X_1(\omega), X_2(\omega)$ is made to either observe $X_3(\omega)$ or to stop with $J(\omega) = 2$, etc.

At each time n, a decision (based on $X_1 \ldots, X_{n-1}$) is first made; if the decision is favorable, then X_n is observed.

At each n, the decision is binary: either proceed to observe X_n (in which case $J \geq n$) or don't observe X_n (in which case $J < n$).

We assume that if X_i is observed but X_{i+1} is not (*i.e.*, $J = i$), then the decision at all $n > i$ is to not observe X_n. (can't restart).

The decision at n is then: choose $J \geq n$ (*i.e.*, make the observation at time n) or choose $J < n$ (*i.e.*, stop at $J = n - 1$ or remain stopped.

The decision at n is then $\mathbb{I}_n$ the indicator rv for $\{J \geq n\}$.

Decisions are rv's?

The decision $\mathbb{I}_n$ at n has been defined as a function of $X_1, \dots, X_{n-1}$, so $\mathbb{I}_n$ is a rv.

The nth decision *rule* could have been chosen at time 0 or time n (or any time before n), but a *rule* chosen at 0 can depend on $X_1, \dots, X_{n-1}$.

Wald's equality ($\sum_{n=1}^{J} X_n = \mathsf{E}[X]\,\mathsf{E}[J]$) is valid for any rule chosen as above (if $\mathsf{E}[J] < \infty$).

No cost extension: the rule at each n can also depend on other rv's, subject to the restriction that $\mathbb{I}_n$ is independent of $X_n, X_{n+1}, \dots,$.

DEF: A stopping time J for a sequence of rv's $X_1, X_2, \ldots,$ is a positive, integer-valued rv such that for each $n \geq 1$, the event $\{J \geq n\}$ is statistically independent of $(X_n, X_{n+1}, \ldots)$.

Note that a stopping time J must be a rv, so J is finite with probability 1, *i.e.*, the observations must eventually stop.

Note also that $J = \mathbb{I}_1 + \mathbb{I}_2 + \cdots$, *i.e.*, if $J = n$, then $\mathbb{I}_1$=1, $\ldots, \mathbb{I}_n$=1 and $\mathbb{I}_{n+1} = 0$, etc.

Wald's equality

If a stopping time is used, especially in gambling, one wants to know what the expected return is, *i.e.*, the return is $S_J = \sum_{n=1}^{J} X_n$ and the expected return is

$$\mathsf{E}[S_J] = \mathsf{E}\left[\sum_{n=1}^{J} X_n\right]$$

Wald's Equality: Let $\{X_n; n \geq 1\}$ be IID rv's, each of mean $\overline{X}$. If J is a stopping time for $\{X_n; n \geq 1\}$ and $\mathsf{E}[J] < \infty$, then

$$\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J]$$

$$\mathsf{E}\left[S_J\right] = \overline{X}\mathsf{E}\left[J\right]$$

Wald's equality essentially says that gambling schemes relying on when to stop are baloney in terms of expected return.

Proof:

$$S_J = \sum_{n=1}^{J} X_n = \sum_{n=1}^{\infty} X_n \mathbb{I}_n$$

$$\begin{aligned}
\mathsf{E}\left[S_J\right] &= \mathsf{E}\left[\textstyle\sum_{n=1}^{\infty} X_n \mathbb{I}_n\right] \\
&= \sum_{n=1}^{\infty} \mathsf{E}\left[X_n \mathbb{I}_n\right] = \sum_{n=1}^{\infty} \overline{X}\mathsf{E}\left[\mathbb{I}_n\right] \\
&= \sum_{n=1}^{\infty} \overline{X}\mathsf{P}\left\{J \geq n\right\} = \overline{X}\mathsf{E}\left[J\right]
\end{aligned}$$

Bizarre coin tossing example

Let X_n be IID, equiprobable ± 1.

Let J be first time n at which $X_1 + \cdots + X_n = 1$

Let $p = \mathsf{P}\{J < \infty\}$. Note that $J = 1$ if $X_1 = 1$.

If $X_1 = -1$, then two subsequent events, each of probability p, are required to reach a sum of 1.

$$p = \frac{1}{2} + \frac{1}{2}p^2$$

The only solution is $p = 1$, so J is a rv.

Observe that $S_J = \sum_n X_n \mathbb{I}_n = 1$ W.P.1. Also, if $\mathsf{E}[J] < \infty$, then $\mathsf{E}[S_J] = \overline{X}\mathsf{E}[J] = 0$.

We conclude that $\mathsf{E}[J] = \infty$.

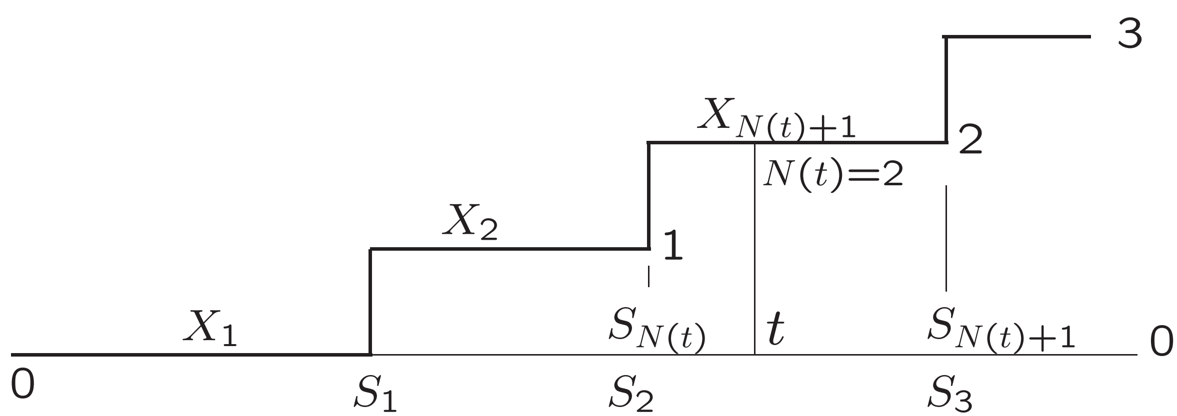
We also conclude that interchanging the expectation and the sum in deriving Wald's equality need not be valid if $\overline{J} = \infty$.

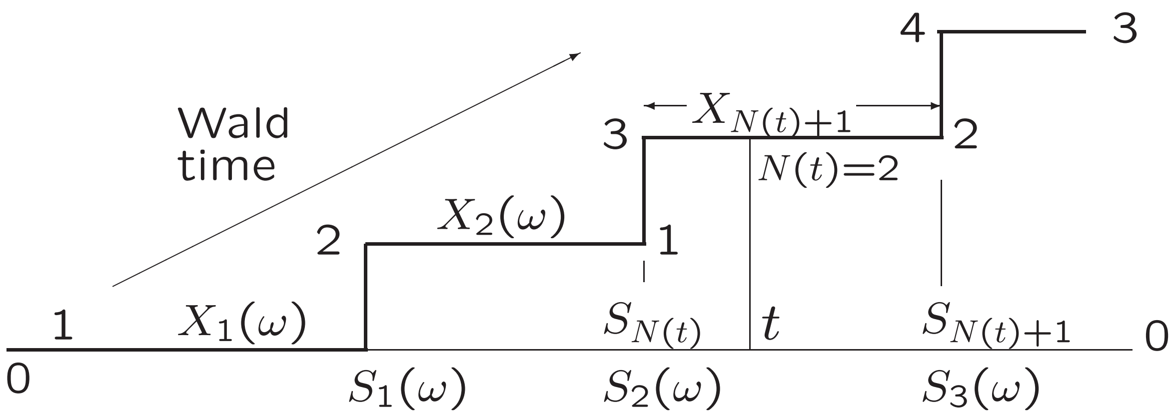
The 'elementary' renewal theorem

Thm: Let $m(t) = \mathsf{E}[N(t)]$ for a renewal process. Then $\lim_{t\to\infty} m(t)/t = 1/\overline{X}$.

We use Wald's equality to prove this. For any given t, suppose we observe $X_1, X_2, \dots$, until the sum exceeds t.

If $N(t) = n$, then $S_{n+1} > t$ so we can stop observing after $n+1$. Thus $N(t) + 1$ is a stopping 'time.'

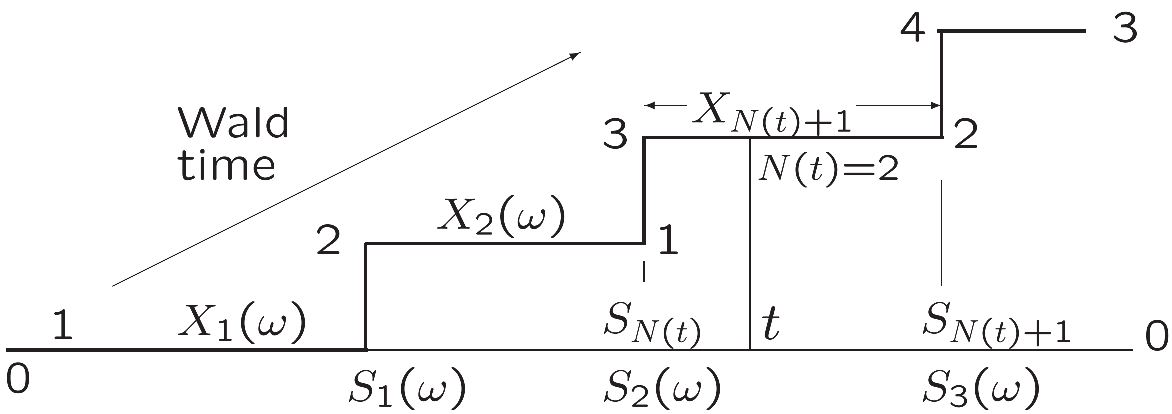




Let $N(t, \omega) = 2$ for example. After observing $X_1(\omega) < t$, the rule at 'Wald time' 2 is to proceed to observe $X_2(\omega)$ ($\mathbb{I}_2 = 1$).

At 'Wald time' 3, based on $S_2(\omega) < t$, proceed to observe $X_3(\omega)$.

At 'Wald time' 4, based on $S_3(\omega) < t$, stop with $J(\omega) = 3$.



Note that $N(t)$ is not a stopping 'time.' for $X_1, X_2, \ldots$,. If $N(t) = n$ and thus $S_n < t$, we can not tell from $X_1, \ldots, X_n$ that $N(t) = n$.

Wald equality (with $N(t) + 1$ as stopping 'time') says

$$\mathsf{E}\left[S_{N(t)+1}\right] = \overline{X}\mathsf{E}\left[N(t) + 1\right]$$

$$\begin{aligned}
\mathsf{E}\left[S_{N(t)+1}\right] &= \overline{X}\mathsf{E}\left[N(t)+1\right] \\
&= \overline{X}(m(t)+1) \\
t &\leq \overline{X}(m(t)+1) \\
\frac{m(t)}{t} &\geq \frac{1}{\overline{X}} - \frac{1}{t}
\end{aligned}$$

We also need an upper bound on $m(t)/t$. This follows from truncating each X_n to $\widetilde{X}_n = \min(X_n, b)$. In the truncated process, $\widetilde{S}_{\tilde{N}(t)+1} \leq t+b$ and $\widetilde{N}(t) \geq N(t)$ so

$$t \geq -b + \mathsf{E}\left[\widetilde{X}\right](m(t)+1)$$

$$\frac{m(t)}{t} \leq \frac{1}{\mathsf{E}\left[\widetilde{X}\right]} - \frac{1}{t} + \frac{b}{t} \quad \rightarrow \frac{1}{\overline{X}} \quad \text{as } t \rightarrow \infty, b = \sqrt{t}$$

Blackwell's theorem

Blackwell's theorem essentially says that the expected renewal rate for large t is $1/\overline{X}$.

It cannot quite say this, since if X is discrete, then S_n is discrete for all n. Thus suggests that $m(t) = \mathsf{E}[N(t)]$ does not have a derivative.

Fundamentally, there are two kinds of distribution funtions — arithmetic and non-arithmetic.

A rv X has an arithmetic distribution if its set of possible sample values are integer multiples of some number. The largest such number is the *span* of the distribution.

If X is arithmetic with span $d > 0$, then every S_n must be arithmetic with a span either d or an integer multiple of d.

Thus $N(t)$ can increase only at multiples of d.

For a non-arithmetic discrete distribution (example: $f_X(1)=1/2$, $f_X(\pi)=1/2$), the points at which $N(t)$ can increase become dense as $t \to \infty$.

Blackwell's thm:

$$\lim_{t\to\infty} [m(t+d) - m(t)] = \frac{d}{\overline{\overline{X}}} \qquad \text{Arith. } X\text{, span } d$$

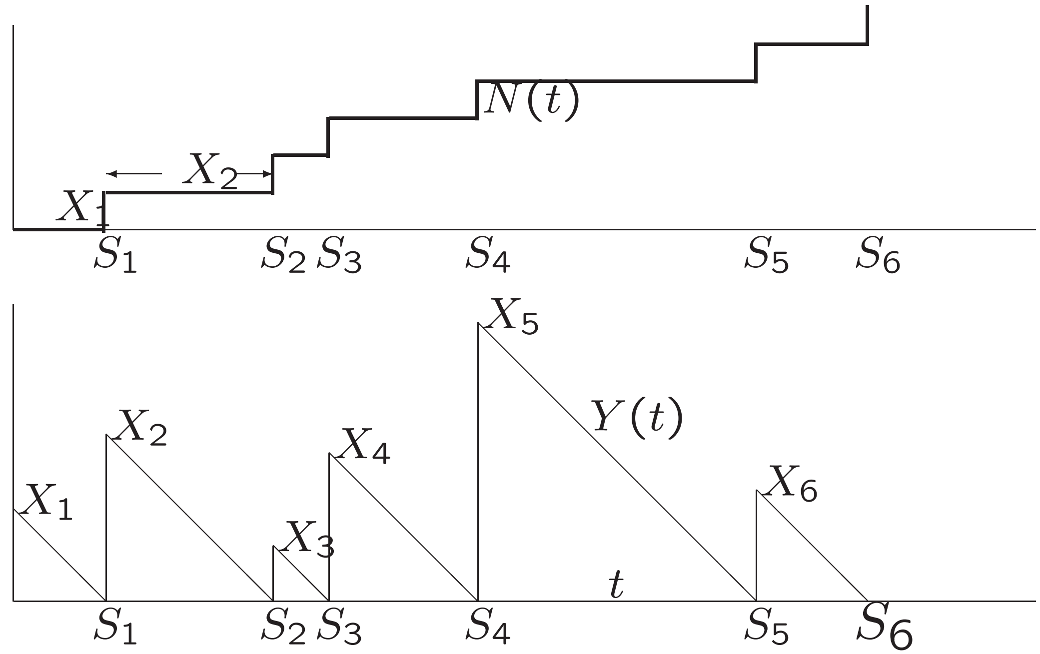
$$\lim_{t\to\infty} [m(t+\delta) - m(t)] = \frac{\delta}{\overline{\overline{X}}} \qquad \text{Non-Arith. } X\text{, any } \delta > 0$$

Time-average Residual life

Def: The residual life $Y(t)$ of a renewal process at time t is the remaining time until the next renewal, *i.e.*, $Y(t) = S_{N(t)+1} - N(t)$.

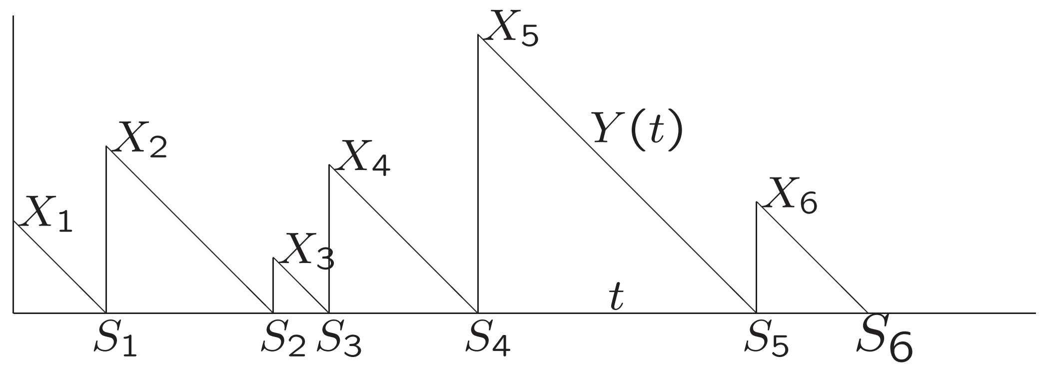
It's how long you have to wait for a bus (if bus arrivals were renewal processes).

The residual life, as a function of t, is a random process, and we can look at its time average value, $t^{-1} \int_0^t Y(\tau)\, d\tau$.



Note that a residual-life sample function is a sequence of isosceles triangles, one starting at each arrival epoch. The time average for a given sample function is

$$\frac{1}{t}\int_0^t y(\tau)d\tau = \frac{1}{2t}\sum_{i=1}^{n(t)} x_i^2 + \frac{1}{t}\int_{\tau=s_{n(t)}}^t y(\tau)d\tau$$



$$\frac{1}{2t}\sum_{n=1}^{N(t)} X_n^2 \le \frac{1}{t}\int_{\tau=0}^{t} Y(\tau)d\tau \le \frac{1}{2t}\sum_{n=1}^{N(t)+1} X_n^2$$

$$\lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{2t} = \lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} X_n^2}{N(t)} \frac{N(t)}{2t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}\left[X\right]} \quad \text{WP1}$$

$$\lim_{t\to\infty} \frac{\int_{\tau=0}^{t} Y(\tau)\,d\tau}{t} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}\left[X\right]} \qquad \text{W.P.1}$$

Time average residual life $\overline{Y}_{ta} = \frac{\mathsf{E}\left[X^2\right]}{2\mathsf{E}[X]}$.

If X is almost deterministic, $\overline{Y}_{ta} \approx \mathsf{E}\left[X\right]/2$.

If X exponential, $\overline{Y}_{ta} = \mathsf{E}\left[X\right]$.