ESD.86
Random Incidence
A Major Source of Selection Bias

Richard C. Larson

February 21, 2007

Examples

- Waiting for a bus at 77 Mass. Avenue.
  - “Clumping”
- Interview passengers disembarking from an airplane.

Doctoral Exam Question

- You arrive at a bus stop where busses arrive according to a Poisson Process with rate λ per unit time.
- Use no-memory property of Poisson processes. Time until next bus arrives has negative exponential density with mean $1/\lambda$.
- Looking backwards, time since last bus was at the bus stop has negative exponential density with mean $1/\lambda$.
- Thus, mean time between buses is $2/\lambda$, not $1/\lambda$. What's wrong here?

Random Incidence: Tending to "Land" in Bigger Gaps

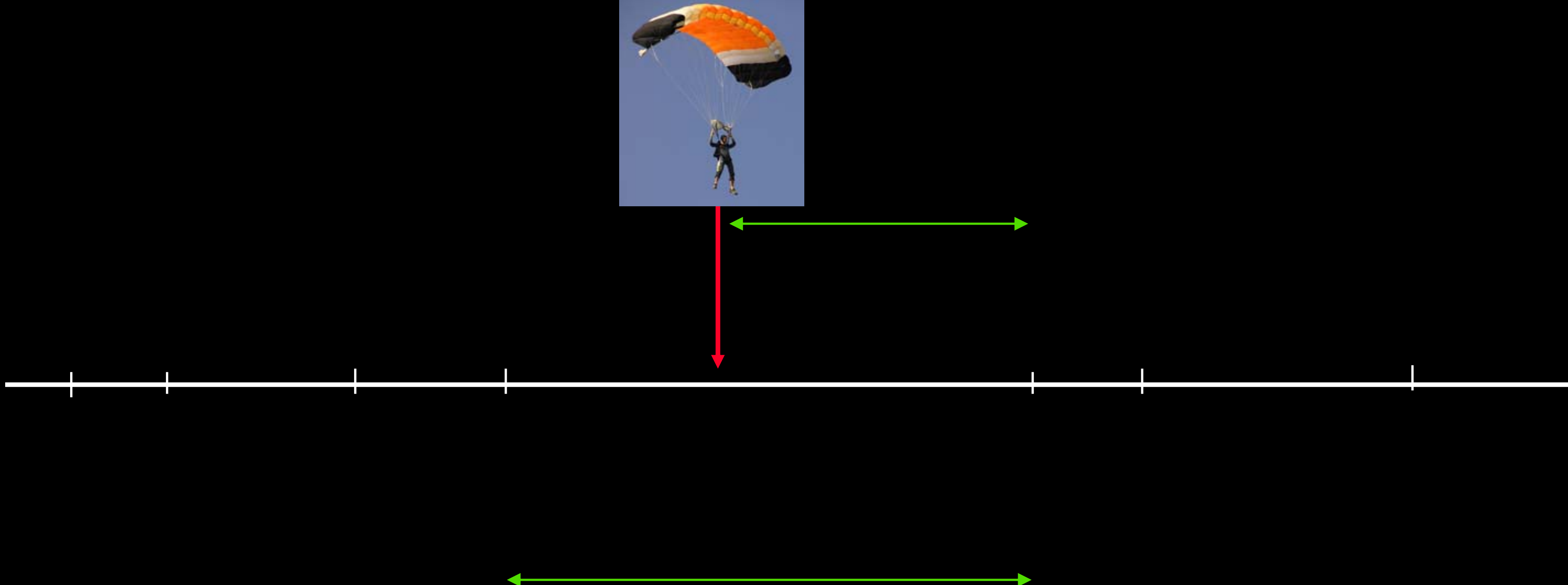
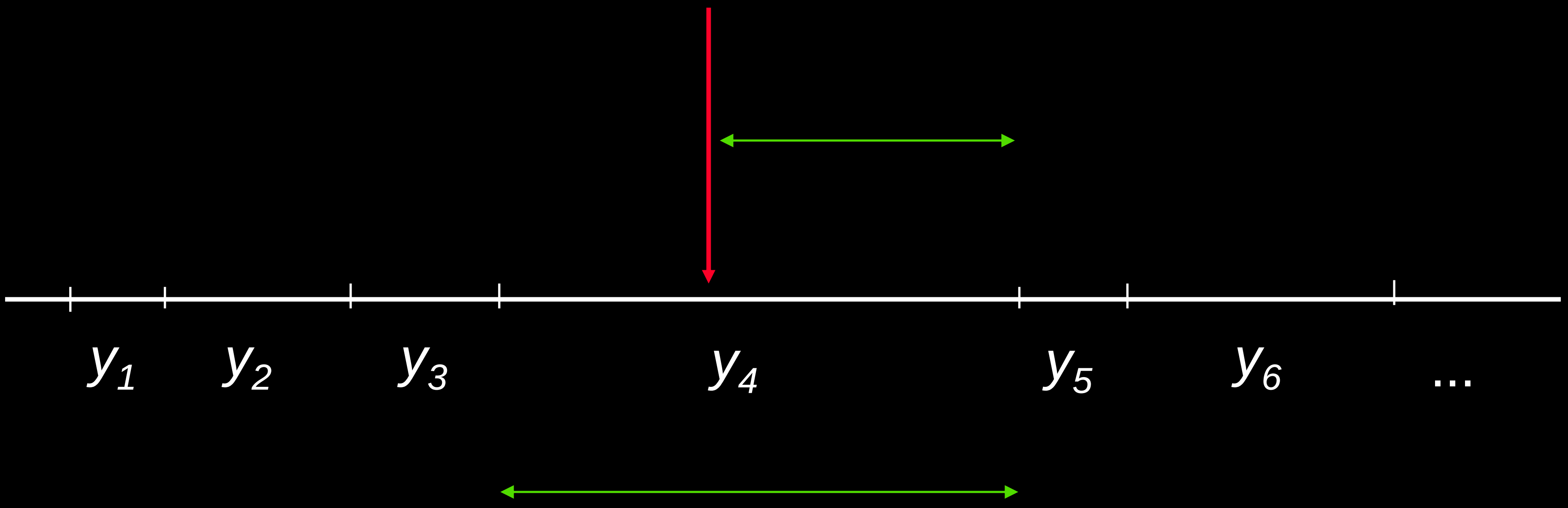
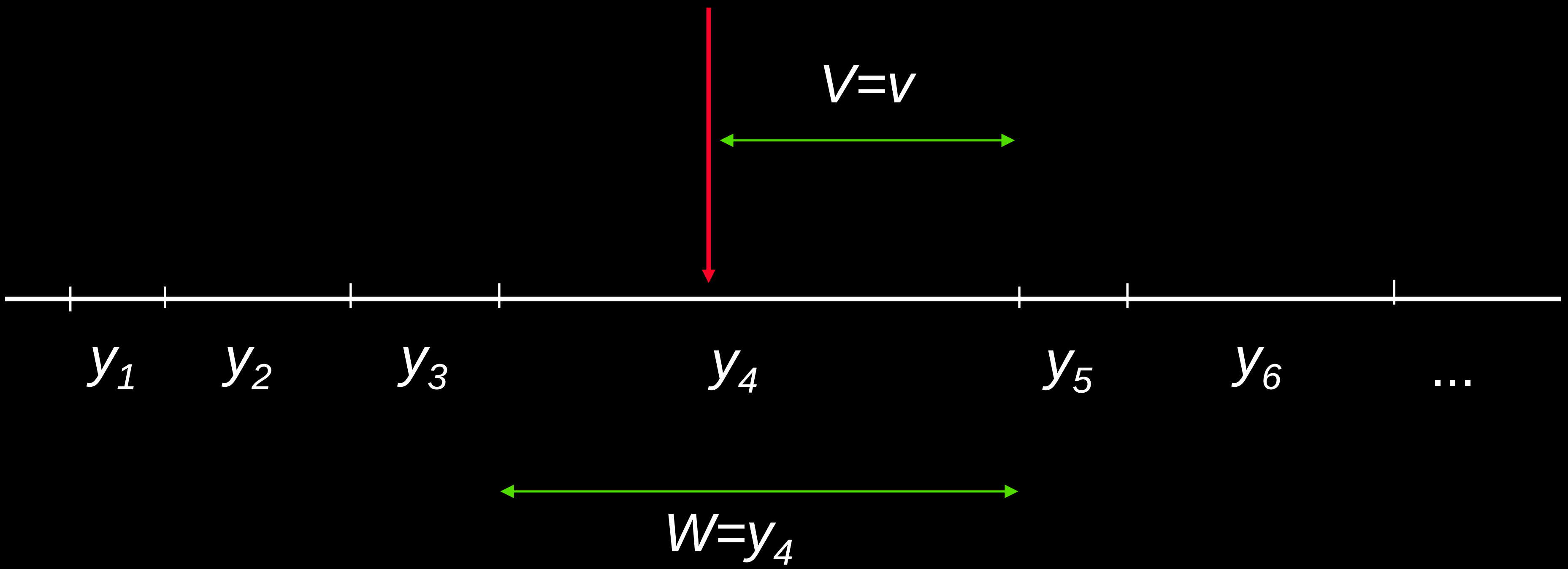


Photo courtesy of Kevin King. http://www.flickr.com/photos/divemasterking2000/541537501/



y_1
y_2
y_3
y_4
y_5
y_6
...

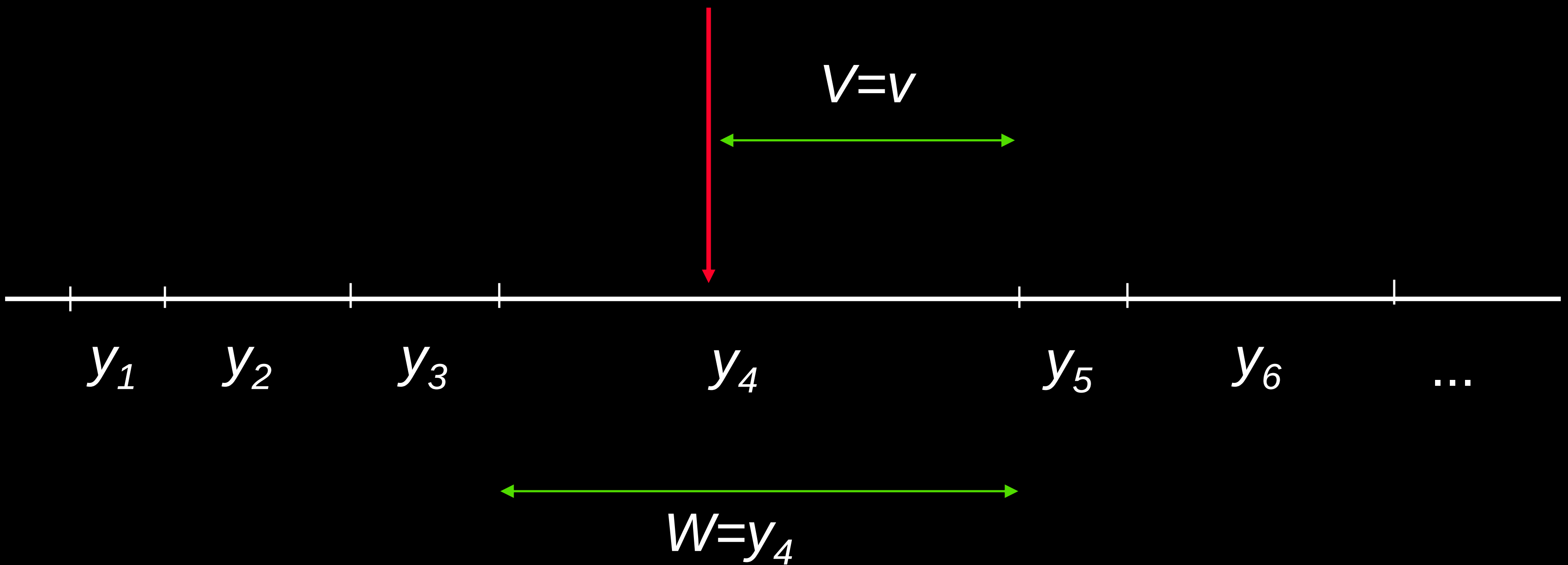


Definitions of the random variables:

Y_i = time interval between the i^{th} and $i + 1^{st}$ arrival event

W = length of the inter-arrival gap in which you fall

V = time remaining in the gap in which you fall



All 3 random variables have probability density functions:

$f_Y(x) = f_{Y_1}(x) = f_{Y_2}(x) = \ldots$

$f_W(w)$

$f_V(y)$

The Inter-Arrival Times

$f_Y(x) = f_{Y_1}(x) = f_{Y_2}(x) = \ldots$

- If the Y_i's are mutually independent then we have a ***renewal process***.
- But the Random Incidence results we are about to obtain do not require that we have a renewal process.

The Gap We Fall Into by Random Incidence

$f_W(w)dw$=P{length of gap is between w and $w+dw$}

$f_W(w)dw$ is proportional to two things:

(1) the relative frequency of gaps *[w, w+dw]*

(2) the length of the gap *w (!!).*

Thus, normalizing so we have a proper pdf,
We can write

$$f_W(w)dw = w f_Y(w) dw / E[Y], \text{ or}$$

$$f_W(w) = w f_Y(w) / E[Y]$$

Time Remaining in the Gap Until Next Arrival

$f_V(y)$

Consider $f_{V|W}(y|w)$

We can argue that $f_{V|W}(y|w) = (1/w)$ for $0 < y < w$.

So we can write

$$f_V(y)dy = dy\int_y^\infty f_{V|W}(y\,|\,w)\,f_W(w)dw$$

$$f_V(y)dy = dy\int_y^\infty (1/w)\frac{wf_Y(w)}{E[Y]}dw$$

$$f_V(y)dy = dy(1 - P\{Y \le y\})/E[Y]$$

Mean Time Until Next Arrival

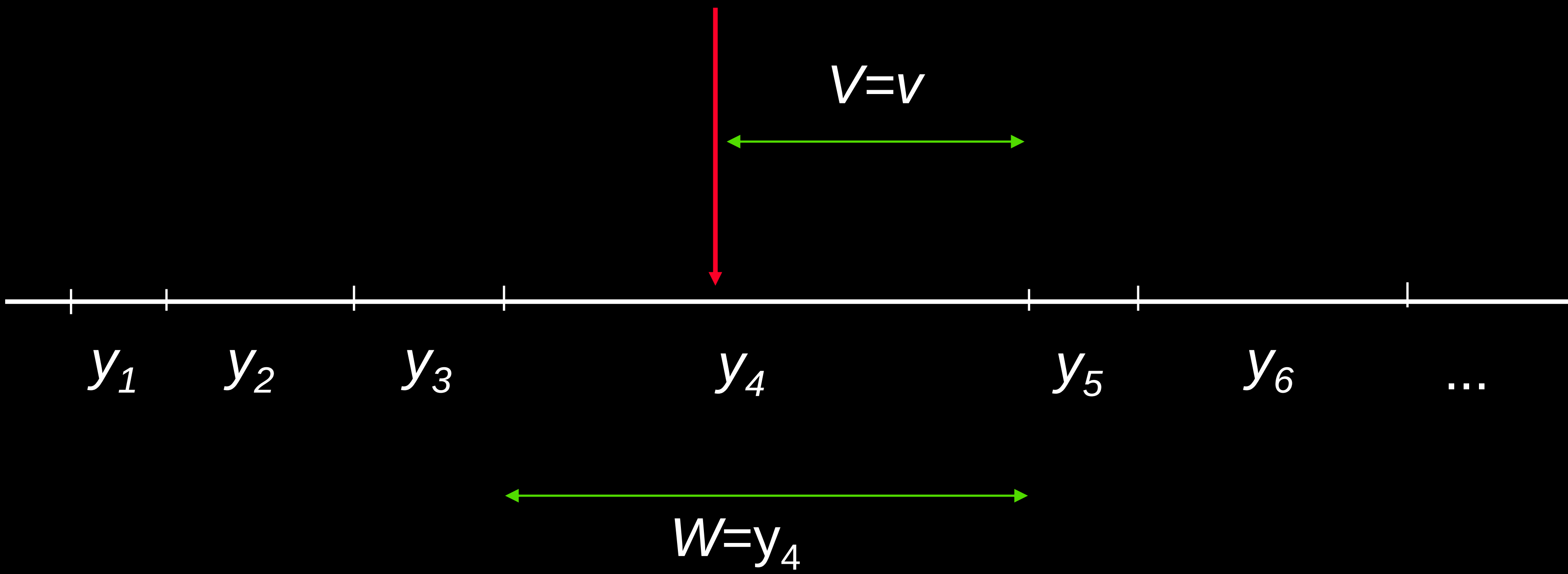
$$E[V] = \int_0^\infty E[V \mid w] f_W(w)\,dw$$

$$E[V] = \int_0^\infty (w/2) \frac{w f_Y(w)}{E[Y]} dw$$

$$E[V] = E[Y^2]/(2E[Y]) = \frac{\sigma_Y^2 + E^2[Y]}{2E[Y]}$$

$$E[V] = E^2[Y] \frac{1 + \sigma_Y^2 / E^2[Y]}{2E[Y]} = E^2[Y] \frac{1 + \eta^2}{2E[Y]},$$

where $\eta \equiv$ coefficient of variation of $Y = \sigma_Y / E[Y]$.



Key result:

$$E[V] = E[Y]^2(1+\eta^2)/(2E[Y])$$

where

$\eta =$ coefficient of variation of the R.V. *Y*

Let’s Visit Several Examples, Including that Bus Stop Doctoral Exam Question!