ESD.86 Feb. 14, 2007

- Broken stick experiment
- D=Min[$X1$, $X2$]
- D=Max[$X1$,$X2$]
- S=$X1$ + $X2$
- Convolution
- Functions of Random Variables

But first, we have a winner!

- **The winning submission for ESD.86, *for most blatant misuse, abuse or misinterpretation of statistics and probability in the media.***
- **Submitted by Roberto Perez-Franco.**

- Original article *New York Times*:

  51% of Women Are Now Living Without Spouse, *New York Times*, January 16, 2007, Section A; Column 1; National Desk; Pg. 1

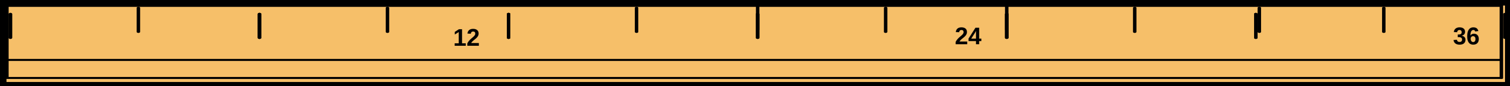
Today: HEARING ON 'WARMING OF PLANET' CANCELED BECAUSE OF ICE STORM

Problem Framing, Formulation and Solution

Break a yardstick in two *random* places

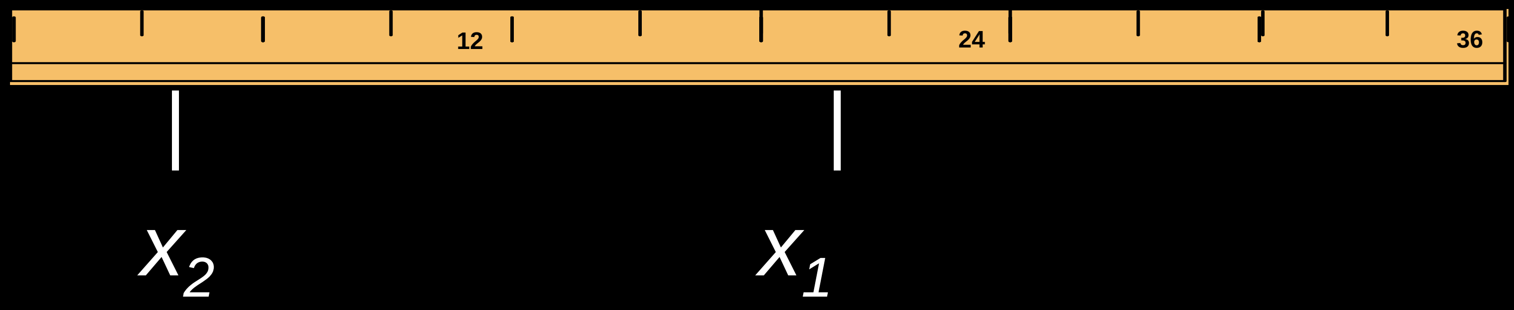
What is the probability that a triangle can be formed with the resulting three stick pieces?

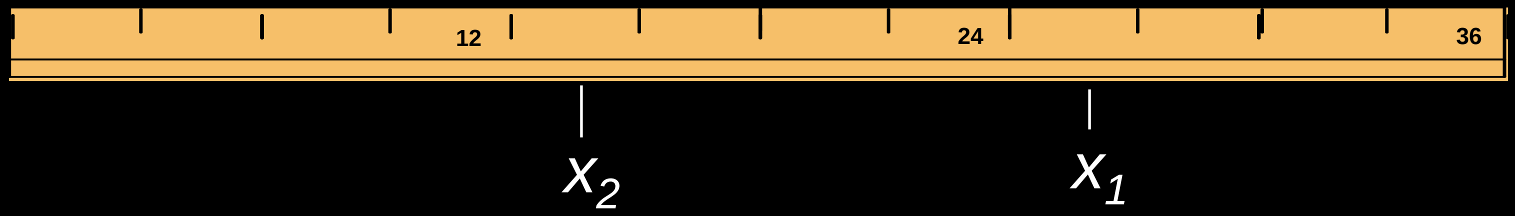
Breaking a Stick



- Mark the stick....

Marking the Results



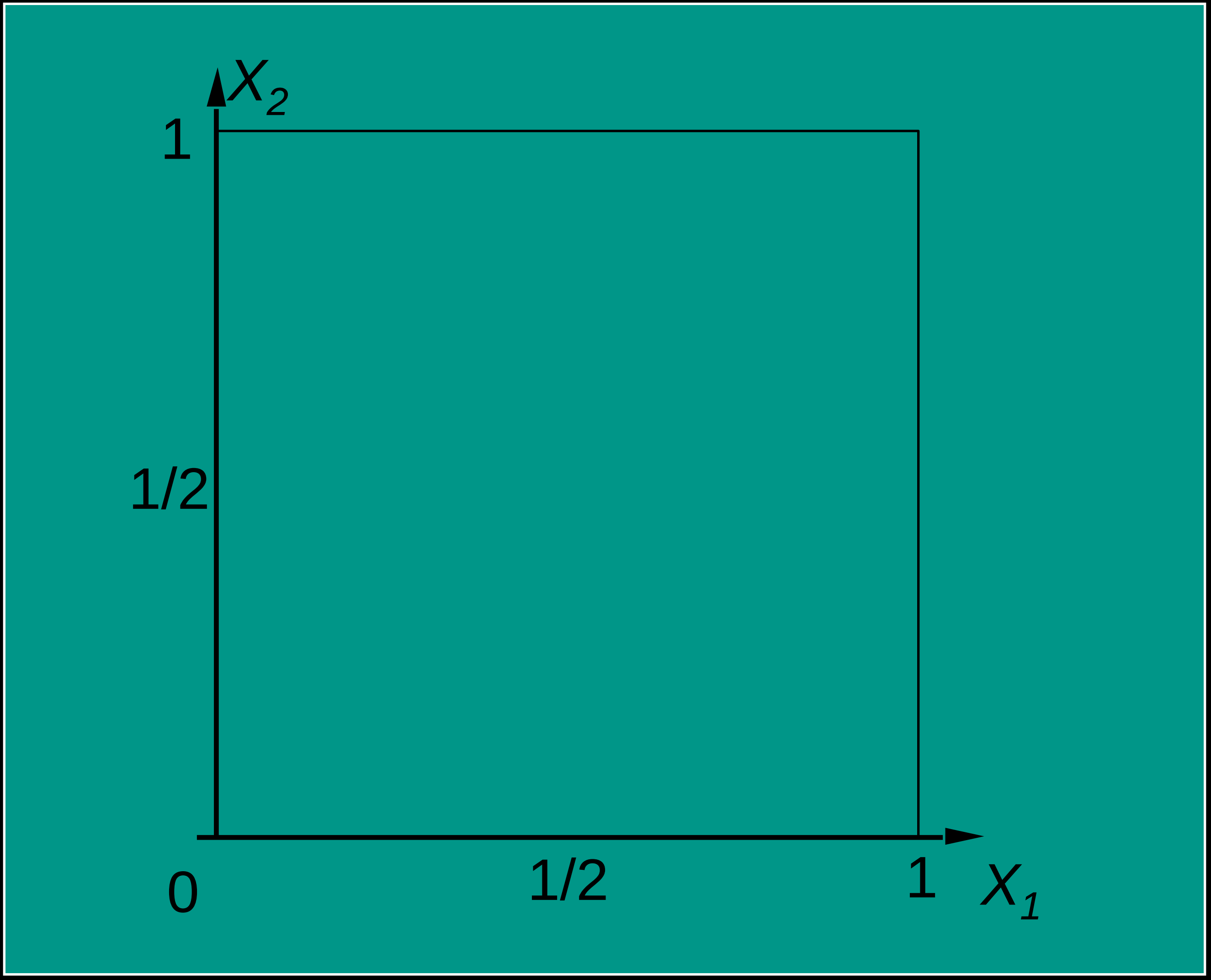


1. Random Variables:

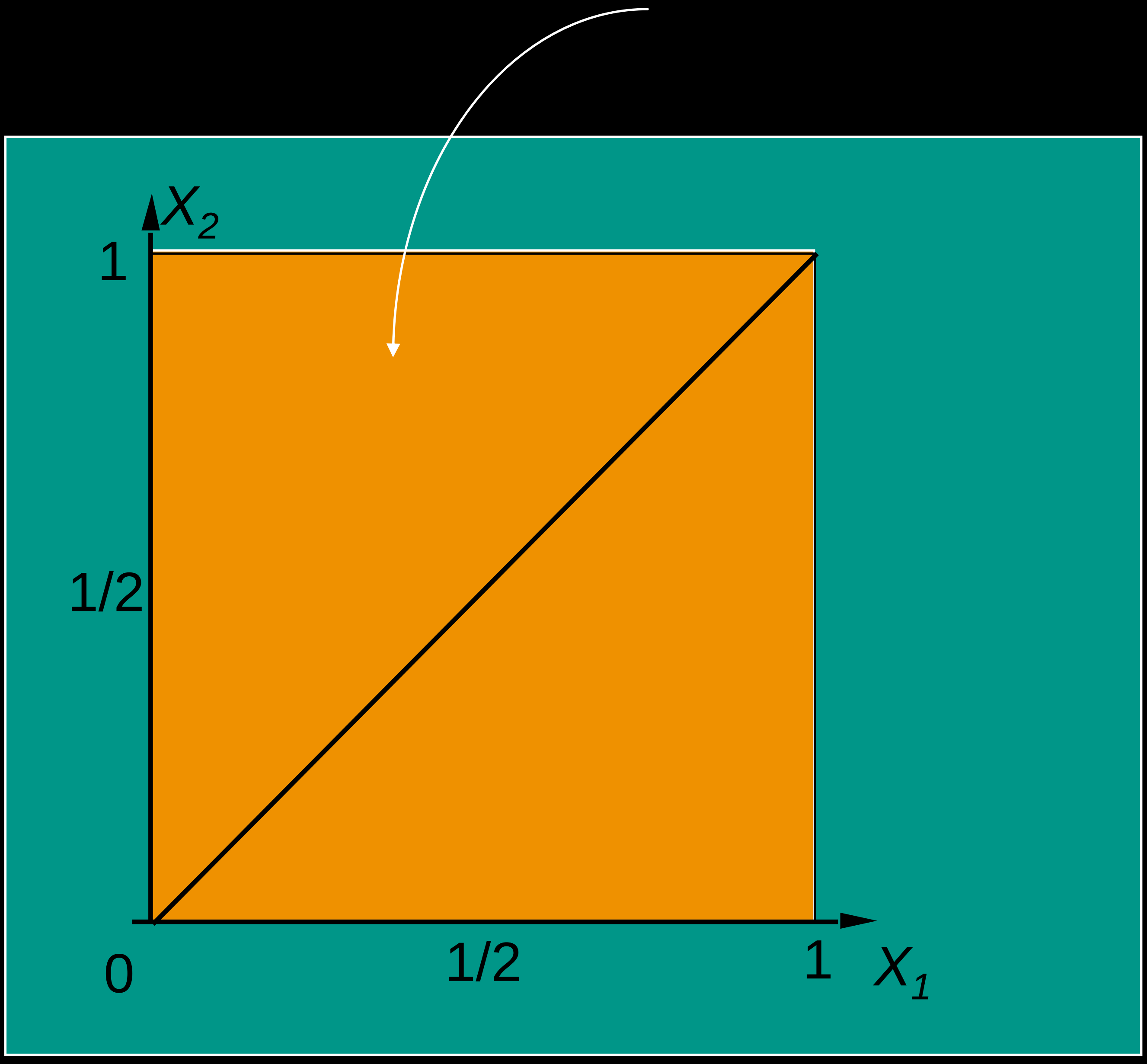
X_1 = location of first mark

X_2 = location of second mark

Step 2: Joint Sample Space

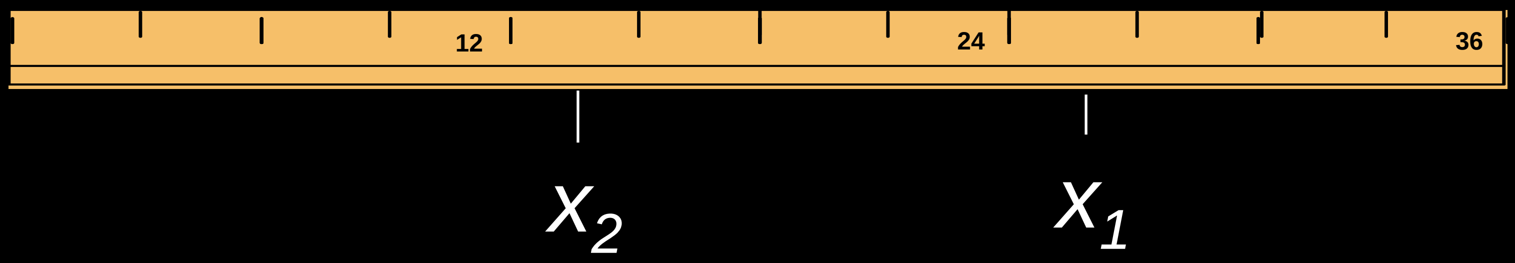


Step 3: Probability Uniform over the Square



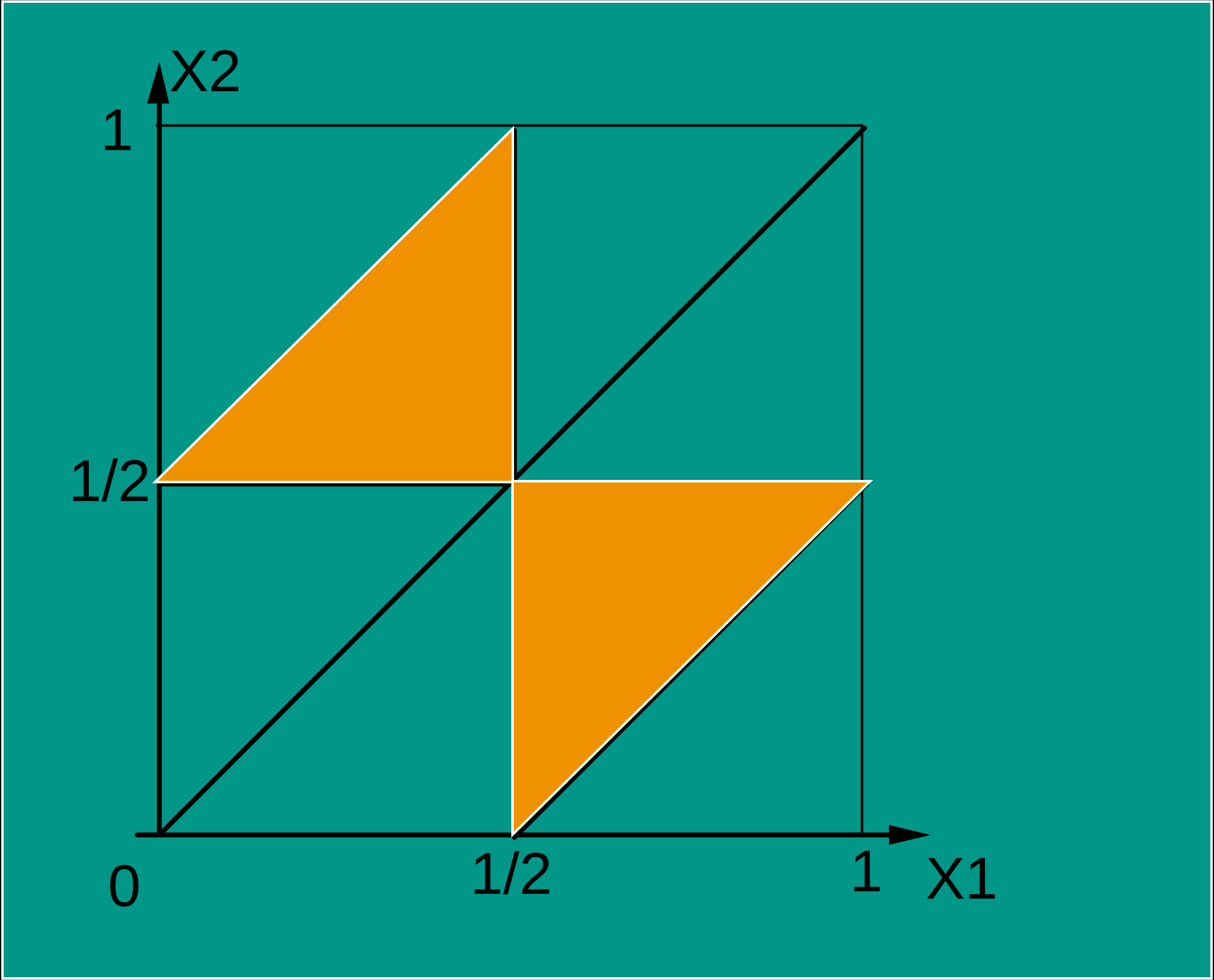
Step 4: Carefully Work Within the Sample Space

- What conditions need to be satisfied so that a triangle can be formed?
- Suppose we consider first the case shown, $x_1 > x_2$



After Step 4, ***HAPPINESS!***

http://web.mit.edu/urban_or_book/www/animated-eg/stick/f1.0.html



Functions of Random Variables

$Y = 3X - 2Z$

4 Steps:

1. Define the Random Variables
2. Identify the joint sample space
3. Determine the probability law over the sample space
4. Carefully work in the sample space to answer any question of interest

4 Steps: Functions of R.V.s

1. Define the Random Variables
2. Identify the joint sample space
3. Determine the probability law over the sample space
4. Carefully work in the sample space to answer any question of interest

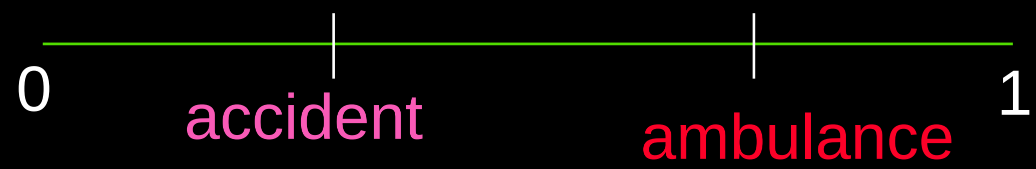
4a. Derive the CDF of the R.V. of interest, working in the original sample space whose probability law you know

4b Take the derivative to obtain the desired PDF

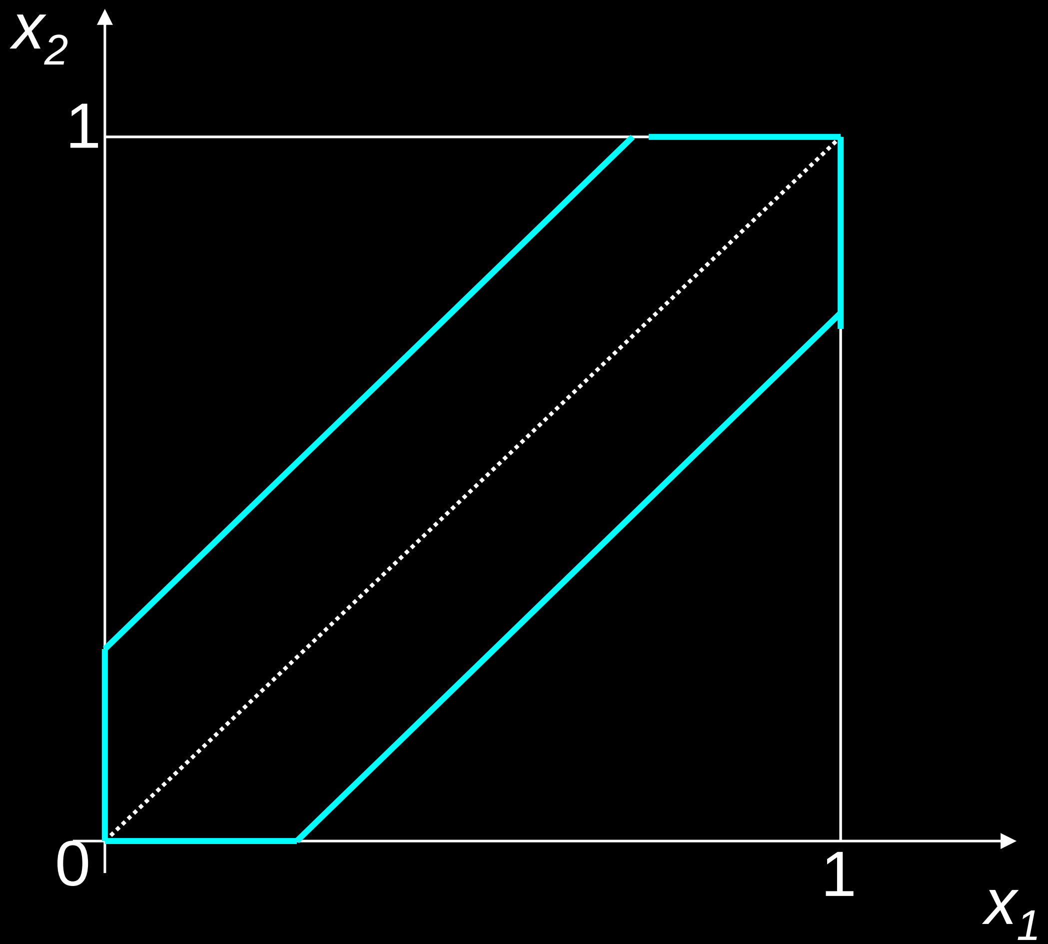
Photos of ambulance and a dispatch center removed due to copyright restrictions.

Response Distance of an Ambulance

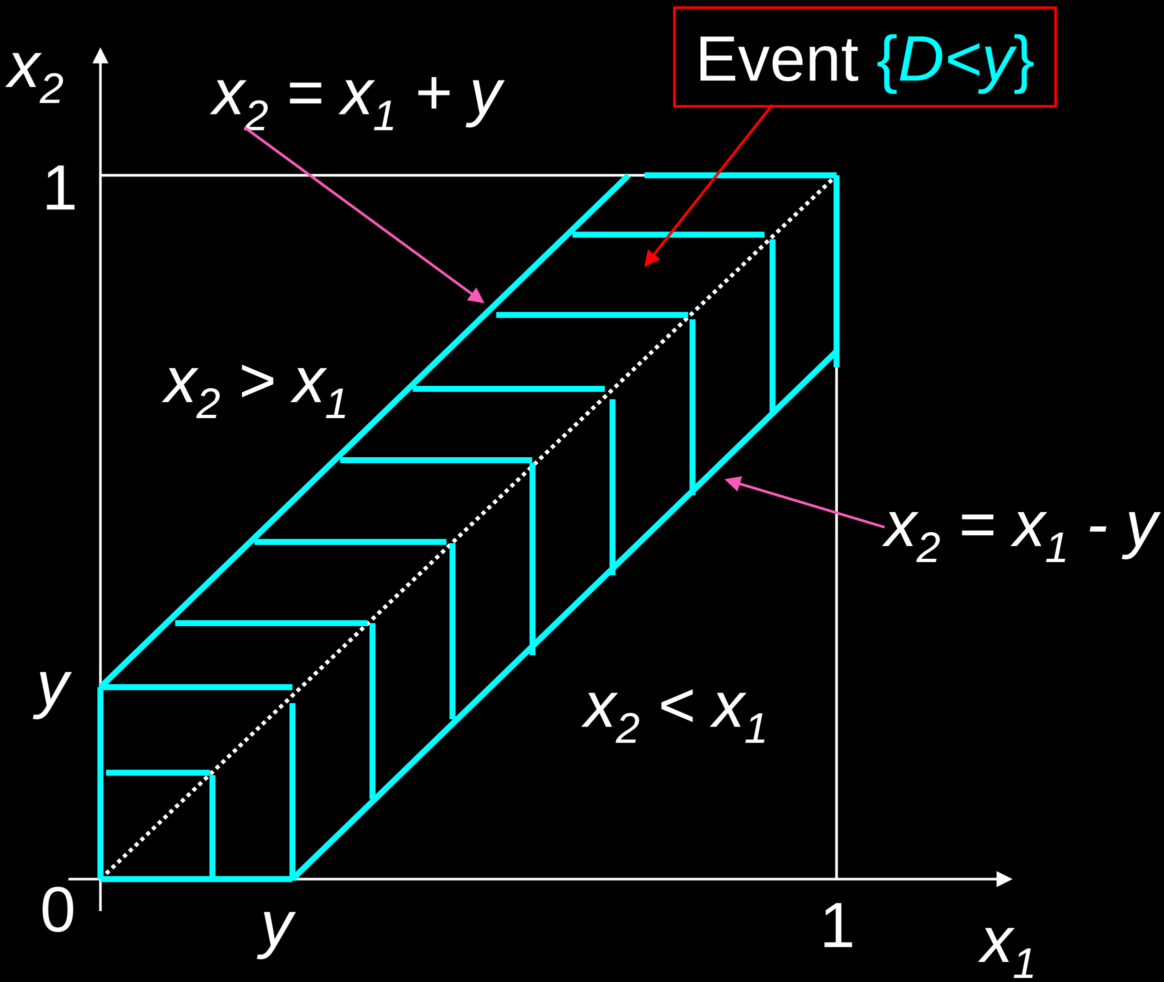
- 1. R.V.;s



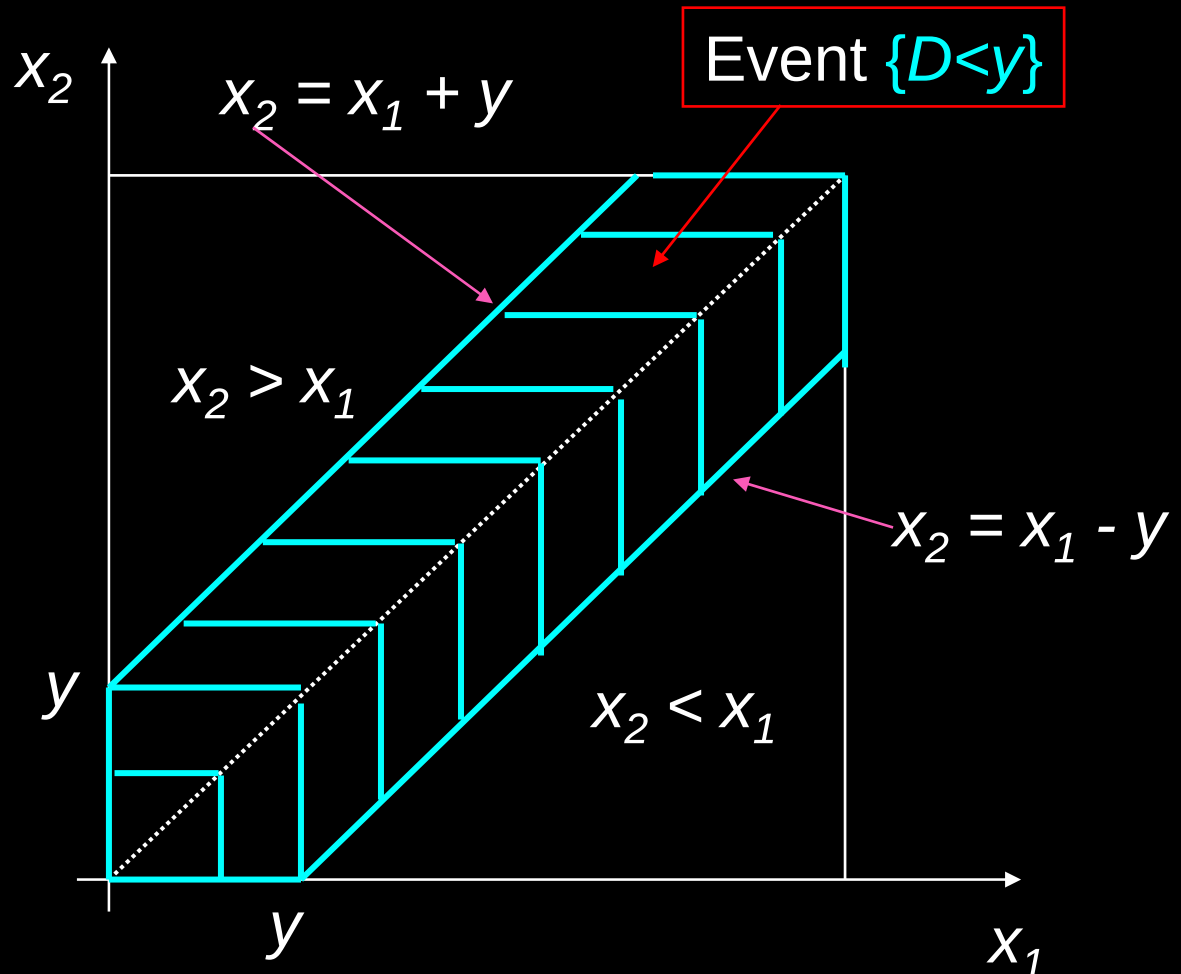
  - X_1 = location of the accident
  - X_2 = location of the ambulance
  - D = response distance = $| X_1 - X_2 |$
- 2. Joint sample space is unit square in X_1 X_2 plane
- 3. PDF over square is uniform



x2
1
0
1
x1

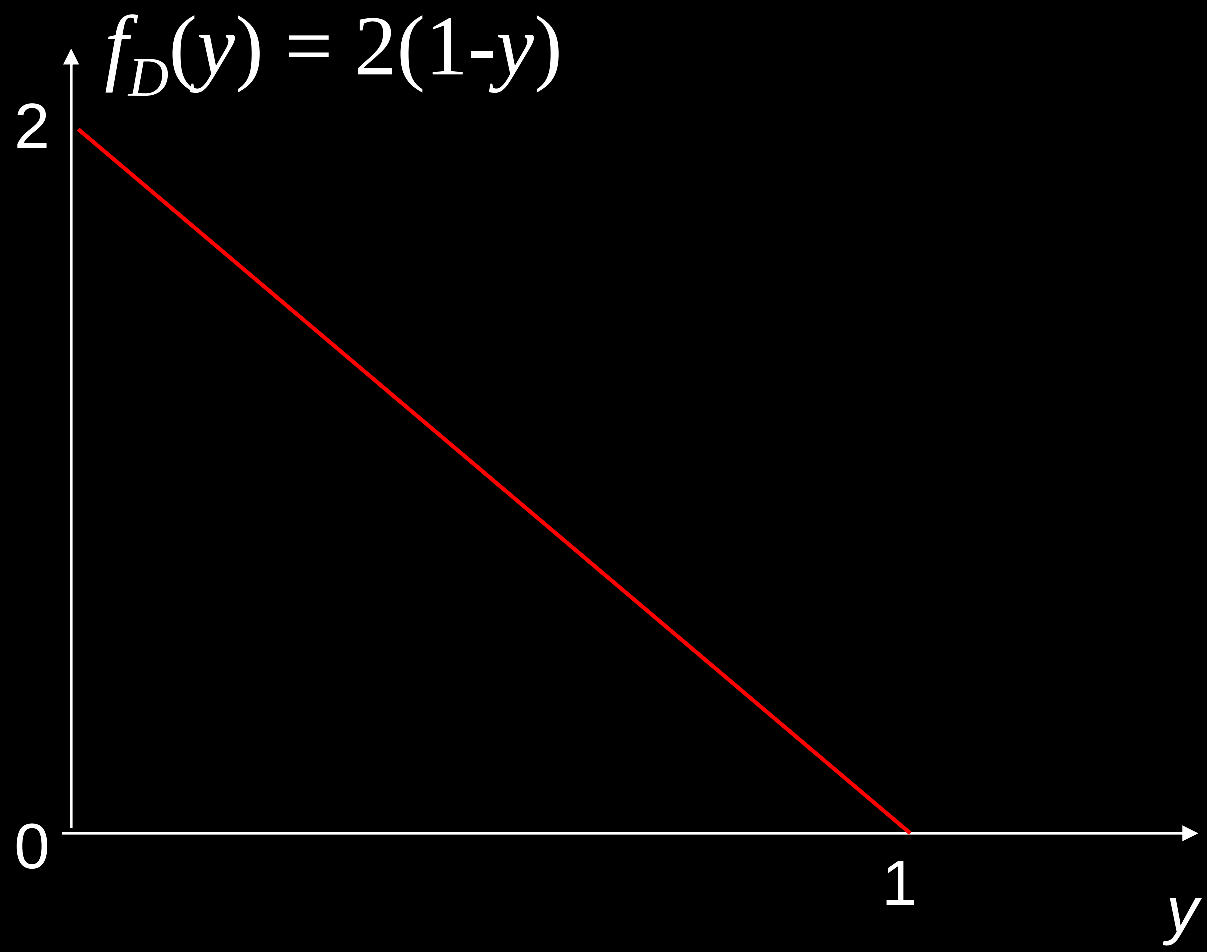


Event {D<y}
x2 = x1 + y
x2 > x1
x2 = x1 - y
x2 < x1
x2
1
y
0
y
1
x1



4.a $F_D(y) = P\{D<y\} = 1 - (1-y)^2, 0<y<1$

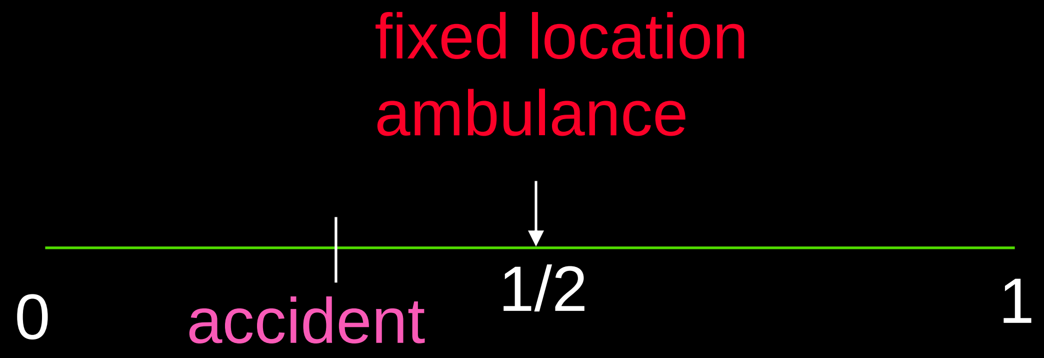
4.b $f_D(y) = 2(1-y), 0<y<1.$

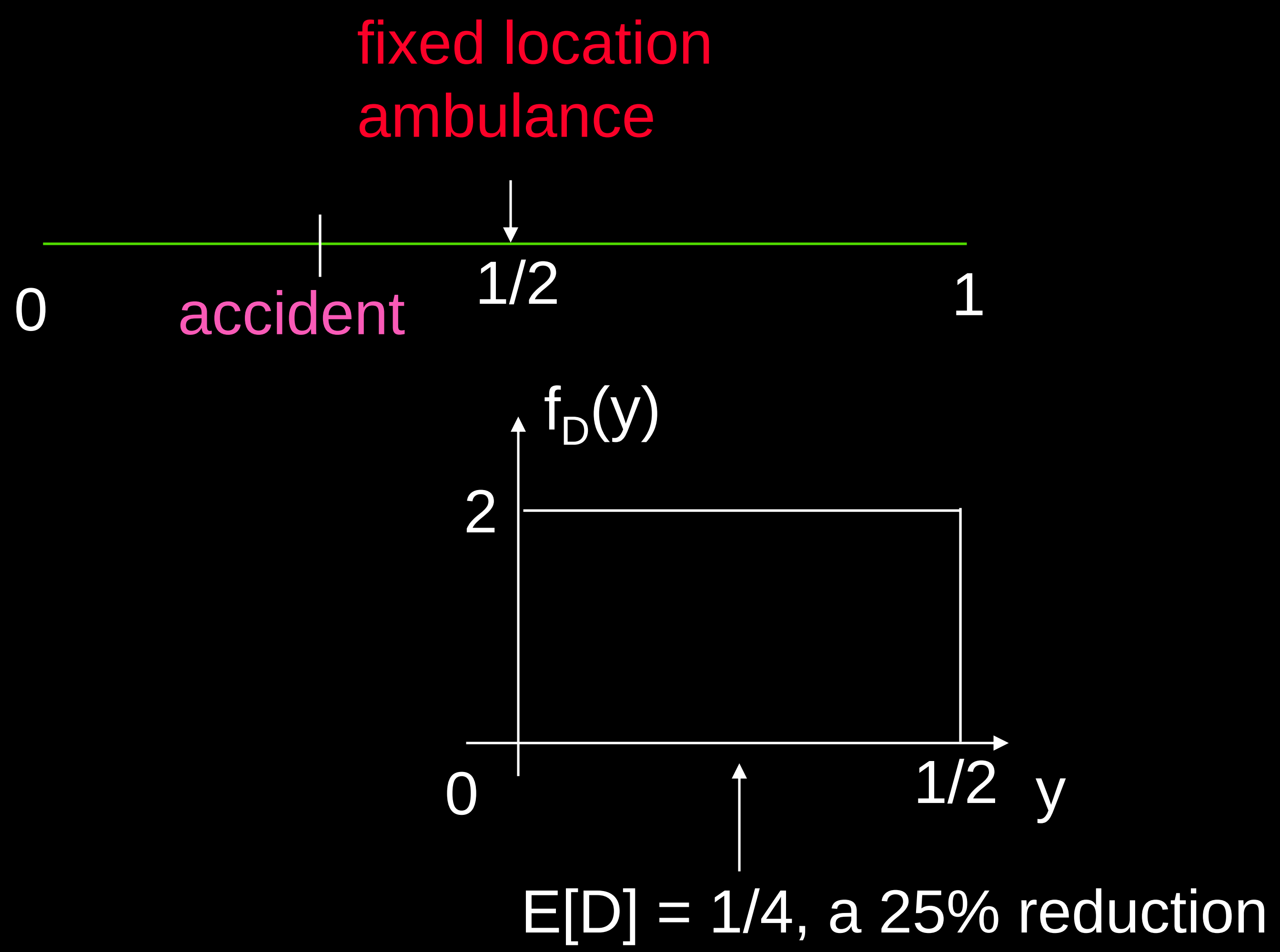


$f_D(y) = 2(1-y)$
2
0
1
y

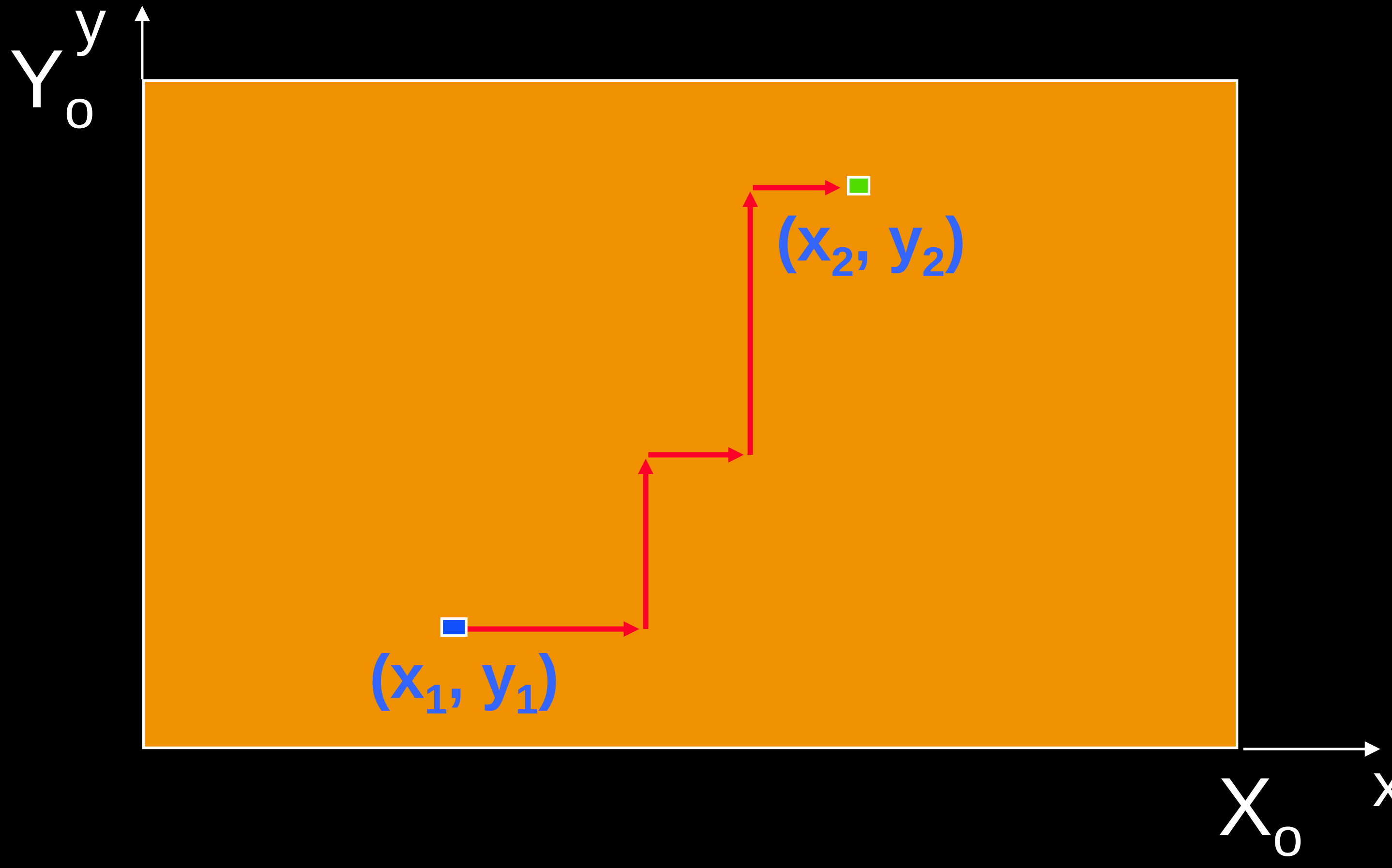
In previous problem, E[D] = 1/3

What if we fix the location of the ambulance at $X_2 = 1/2$?



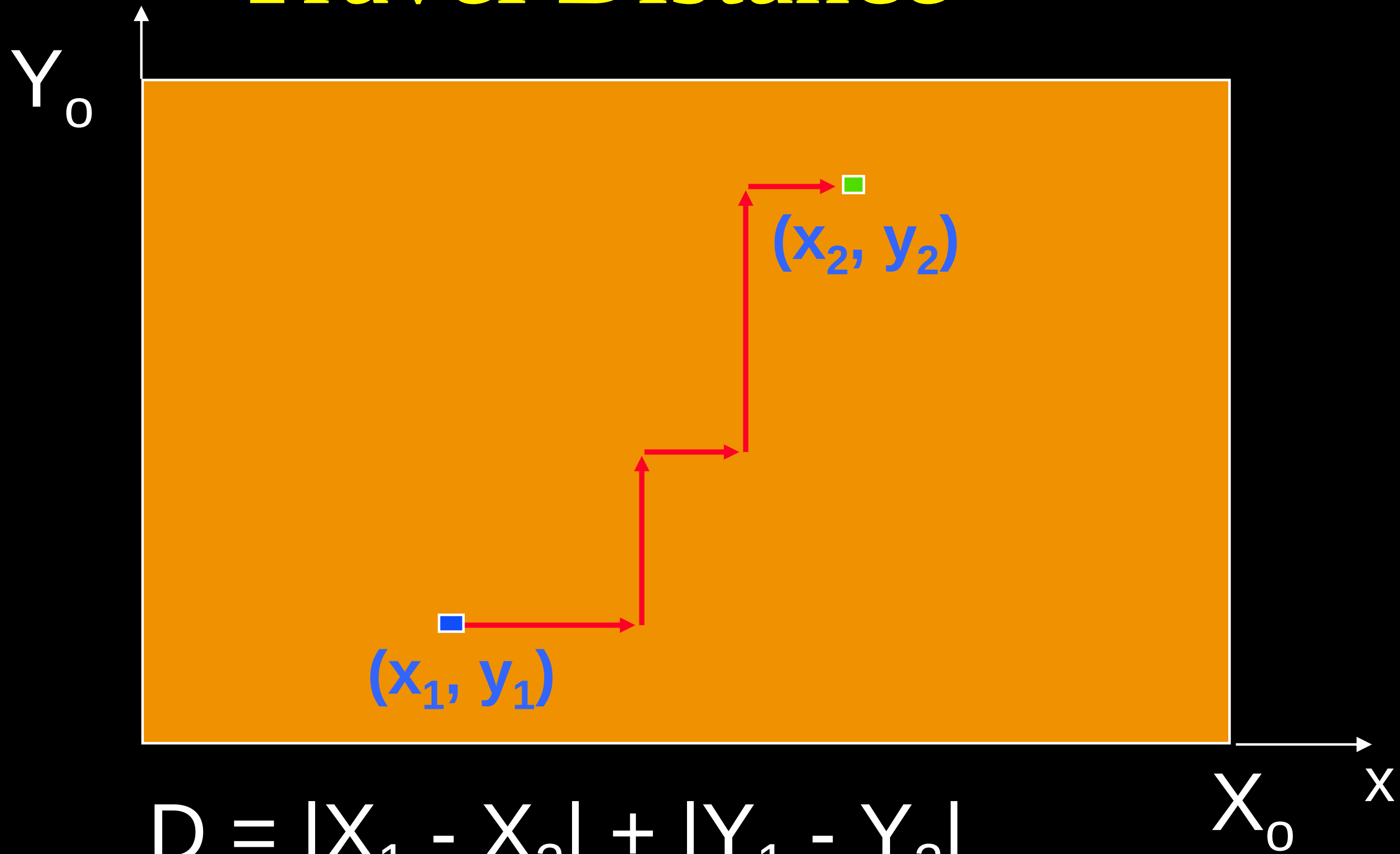

fixed location
ambulance
0
accident
1/2
1
$f_D(y)$
2
0
1/2
y
E[D] = 1/4, a 25% reduction

Rectangular Response Area



$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

Scaling to Get Expected Travel Distance



$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

$$E[D] = E[|X_1 - X_2| + |Y_1 - Y_2|]$$

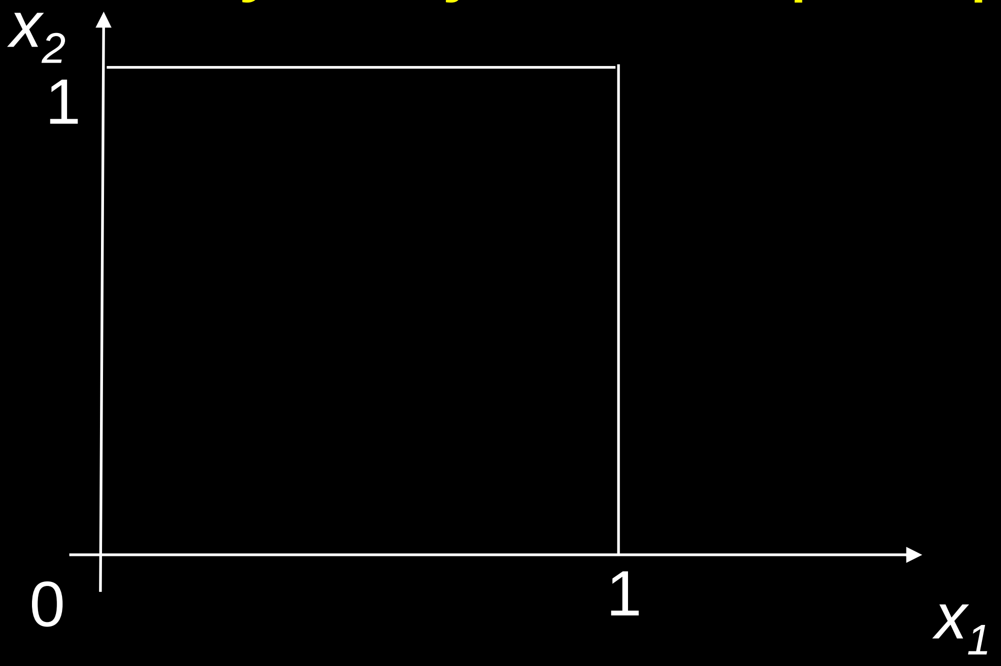
$$E[D] = (1/3)[X_o + Y_o]$$

More Examples of Functions of Random Variables

1. Define the Random Variables

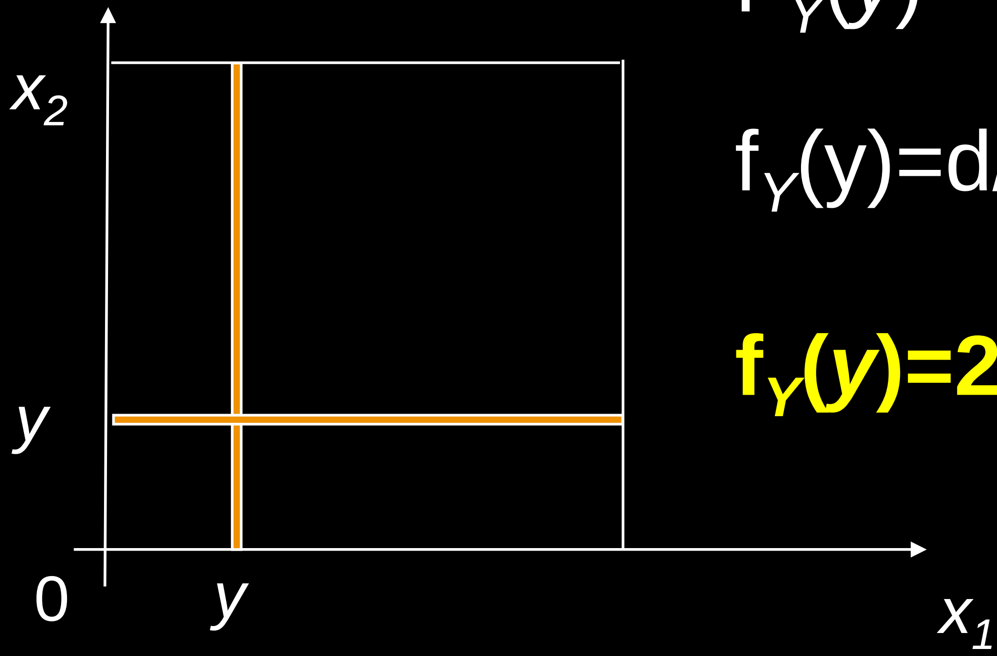
Y=MIN{X_1, X_2}, where X_1 and X_2 are iid uniform over [0,1]

- Identify the joint sample space



3. Determine the probability law over the sample space - uniform

4. Carefully work in the sample space to answer any question of interest.



$F_Y(y) = P\{Y<y\}=1-(1-y)^2$

$f_Y(y)=d/dy\ [F_Y(y)]$

$\mathbf{f_Y(y)=2(1-y),\ 0<y<1}$

4a. Derive the CDF of the R.V. of interest, working in the original sample space whose probability law you know

4b Take the derivative to obtain the desired PDF

Now suppose

Y=MIN{X_1, X_2, X_3, ...X_N}, where X_i are iid uniform over [0,1]

$$F_Y(y) = P\{Y<y\} = 1 - P\{Y>y\}$$

$$F_Y(y) = 1 - (1-y)^N$$

$$\mathbf{f_Y(y) = (d/dy)\, F_Y(y) = N(1-y)^{N-1};\quad N=1,2,\ldots}$$

$$\mathbf{0<y<1}$$

Now suppose

Y=MAX{$X_1, X_2, X_3, \ldots X_N$}, where X_i are iid uniform over [0,1]

$F_Y(y) = P\{Y<y\} = y^N$ *Why?*

$f_Y(y) = Ny^{N-1}$ $N=1,2,\ldots;\ 0<y<1$

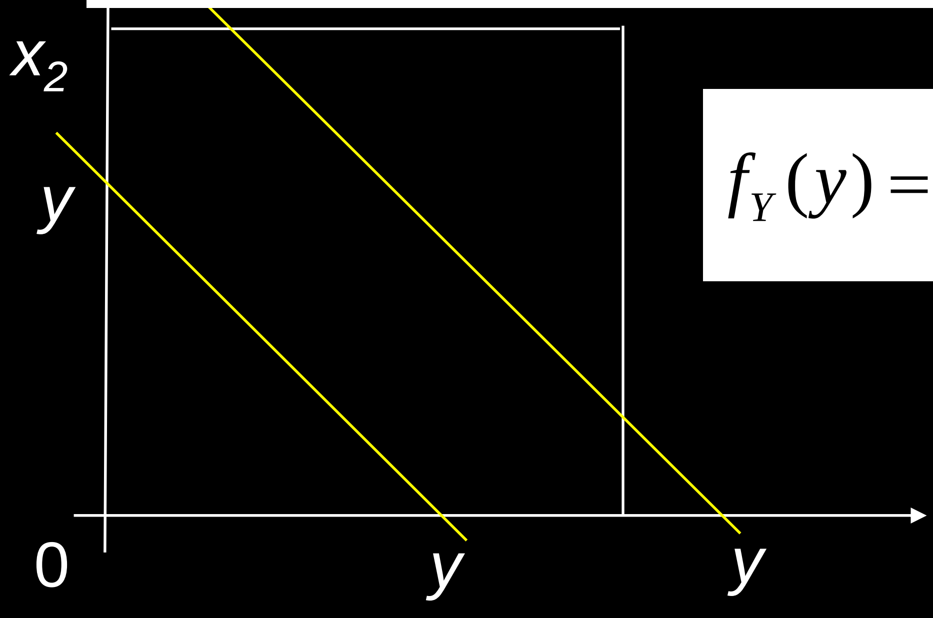
OK, so now we can do Max and Min.

Sums of Random Variables

Now let

$Y=X_1 + X_2$, where X_1 and X_2 are iid uniform over [0,1]

$$F_Y(y) = P\{Y \le y\} = \begin{cases} y^2/2 & 0 \le y \le 1 \\ 1-(2-y)^2/2 & 1 \le y \le 2 \end{cases}$$



$$f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 2-y & 1 \le y \le 2 \end{cases}$$

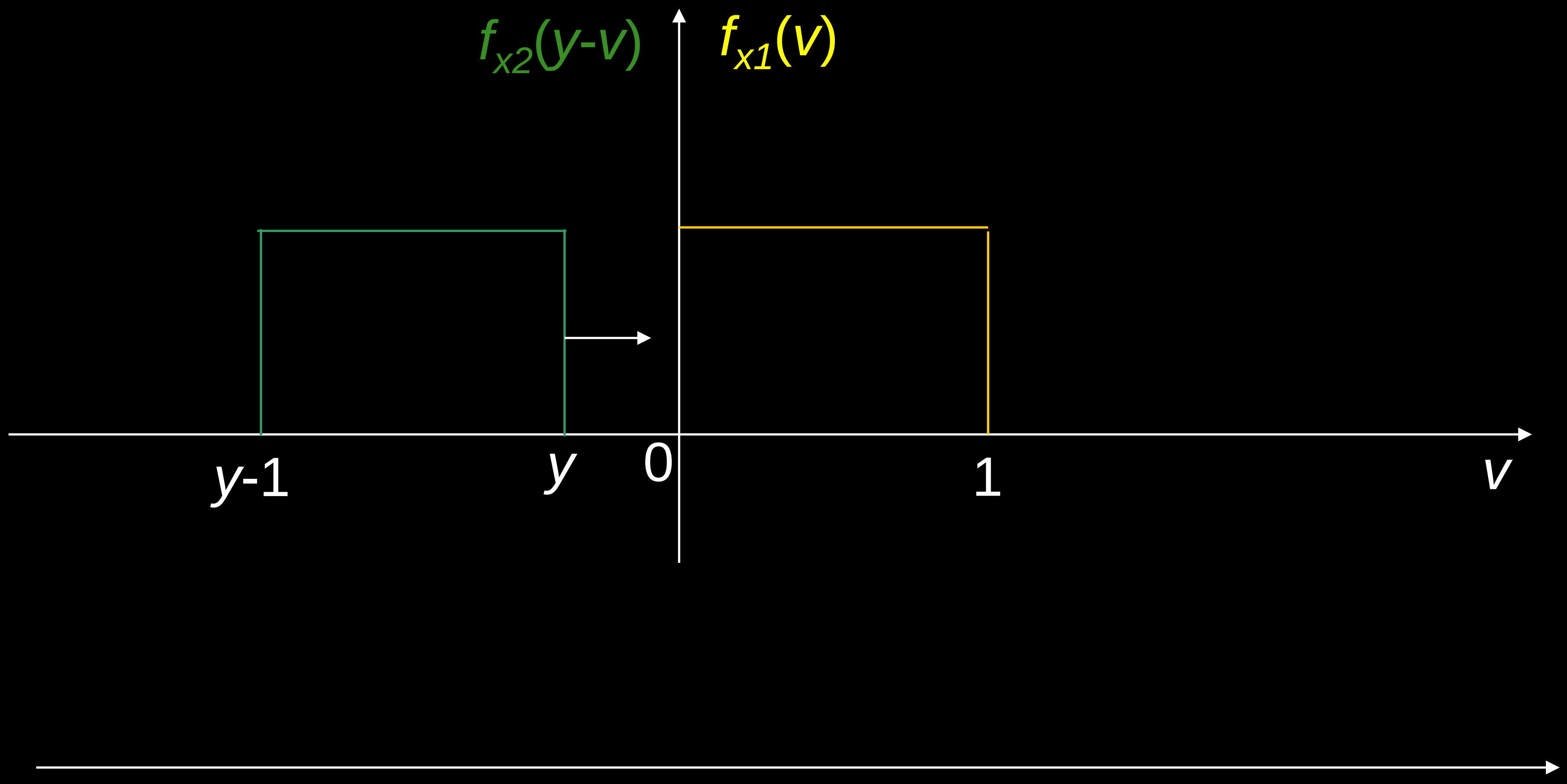
$$f_Y(y)dy = \int_{v=0}^{v=1} f_{x_1}(v) f_{x_2}(y-v)dvdy$$

Convolution

Convolution

$$f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 1-y & 1 \le y \le 2 \end{cases}$$

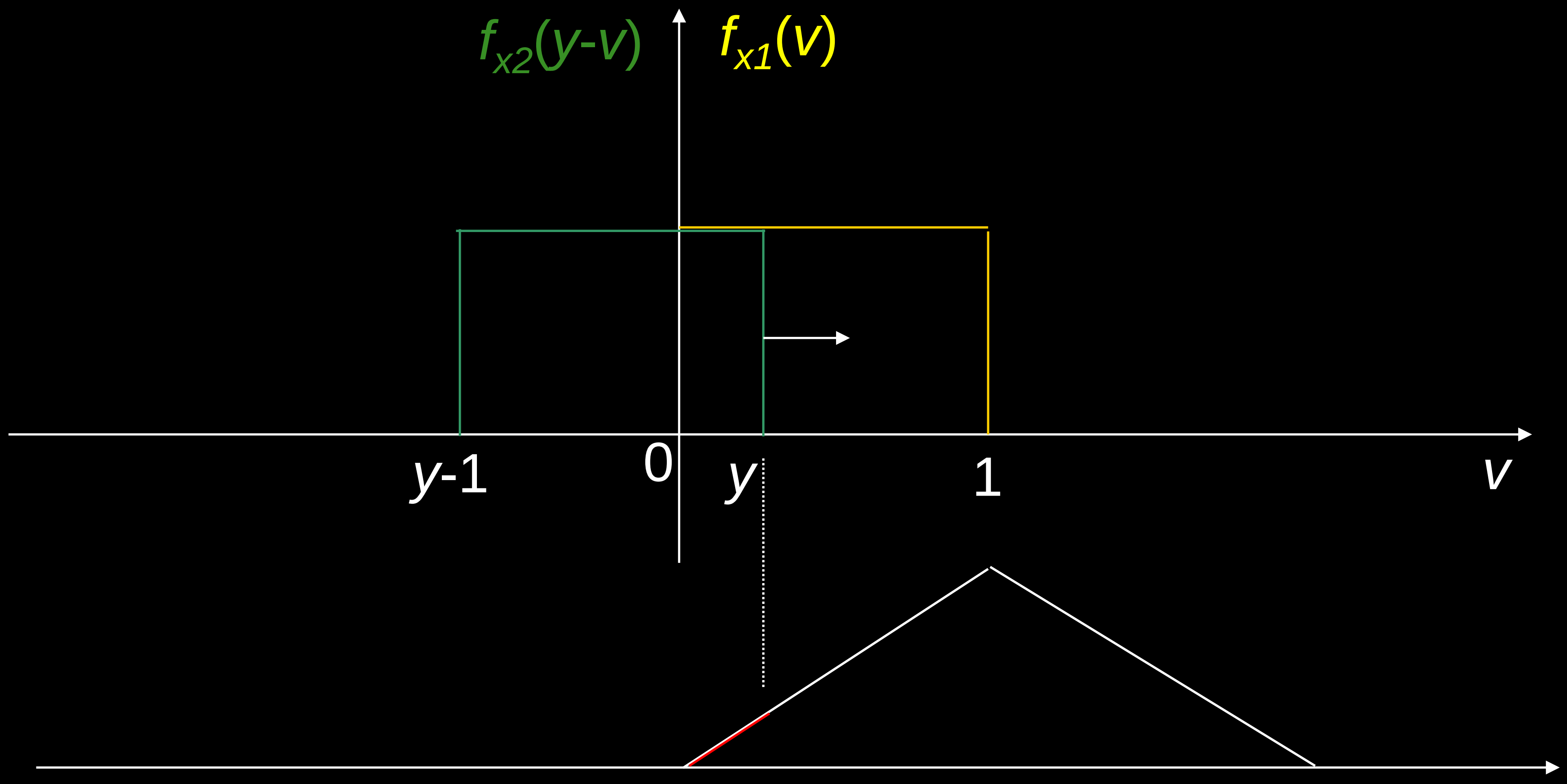
$$f_Y(y)dy = \int_{v=0}^{v=1} f_{x1}(v)\, f_{x2}(y-v)\, dv dy$$



Convolution

$$f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 1 - y & 1 \le y \le 2 \end{cases}$$

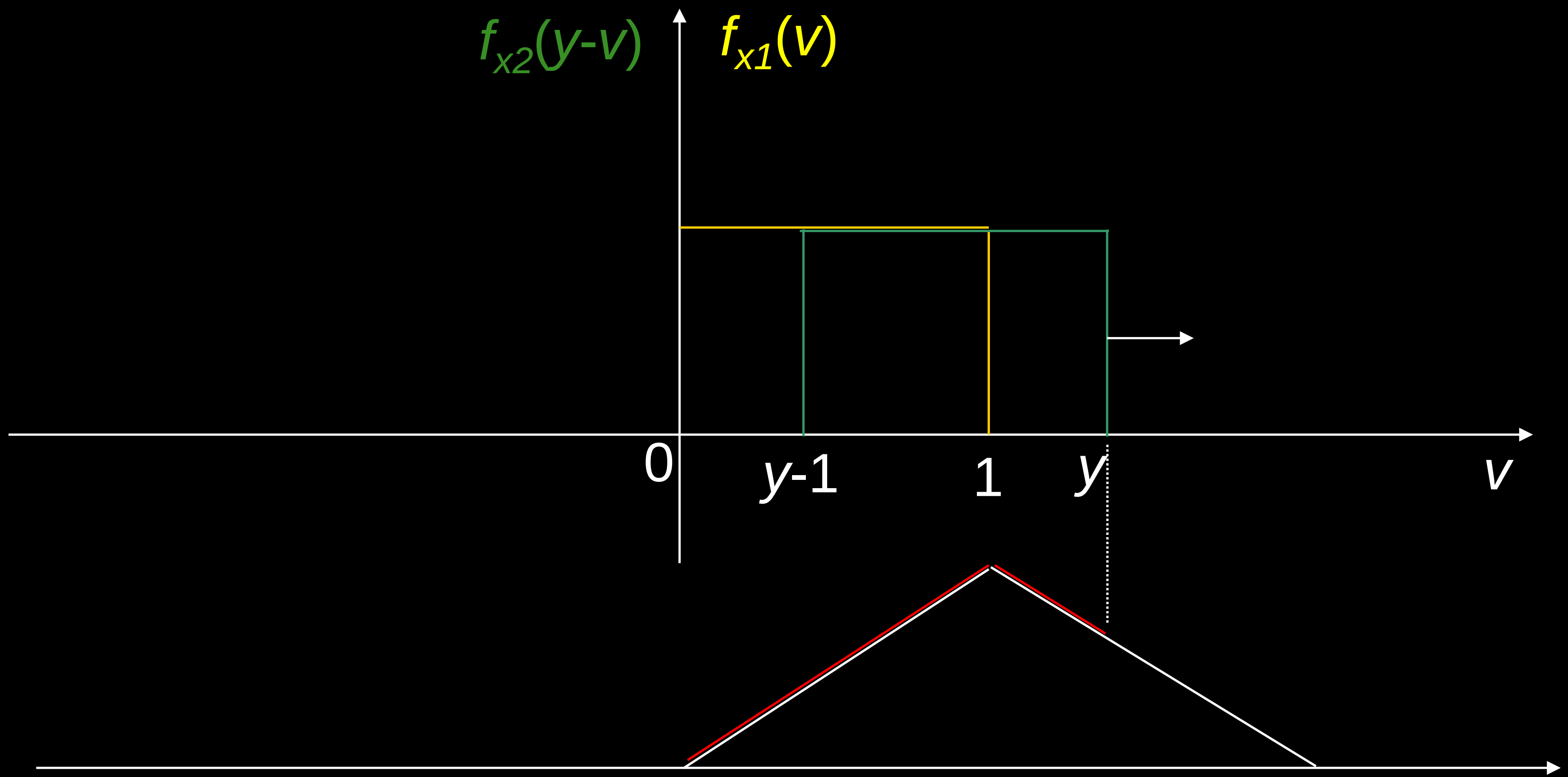
$$f_Y(y)dy = \int_{v=0}^{v=1} f_{x1}(v)\, f_{x2}(y-v)\, dv\, dy$$



$$f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 1 - y & 1 \le y \le 2 \end{cases}$$

Convolution

$$f_Y(y)dy = \int_{v=0}^{v=1} f_{x1}(v)\, f_{x2}(y-v)\, dv\, dy$$



A Quantization Problem

Barges in Action

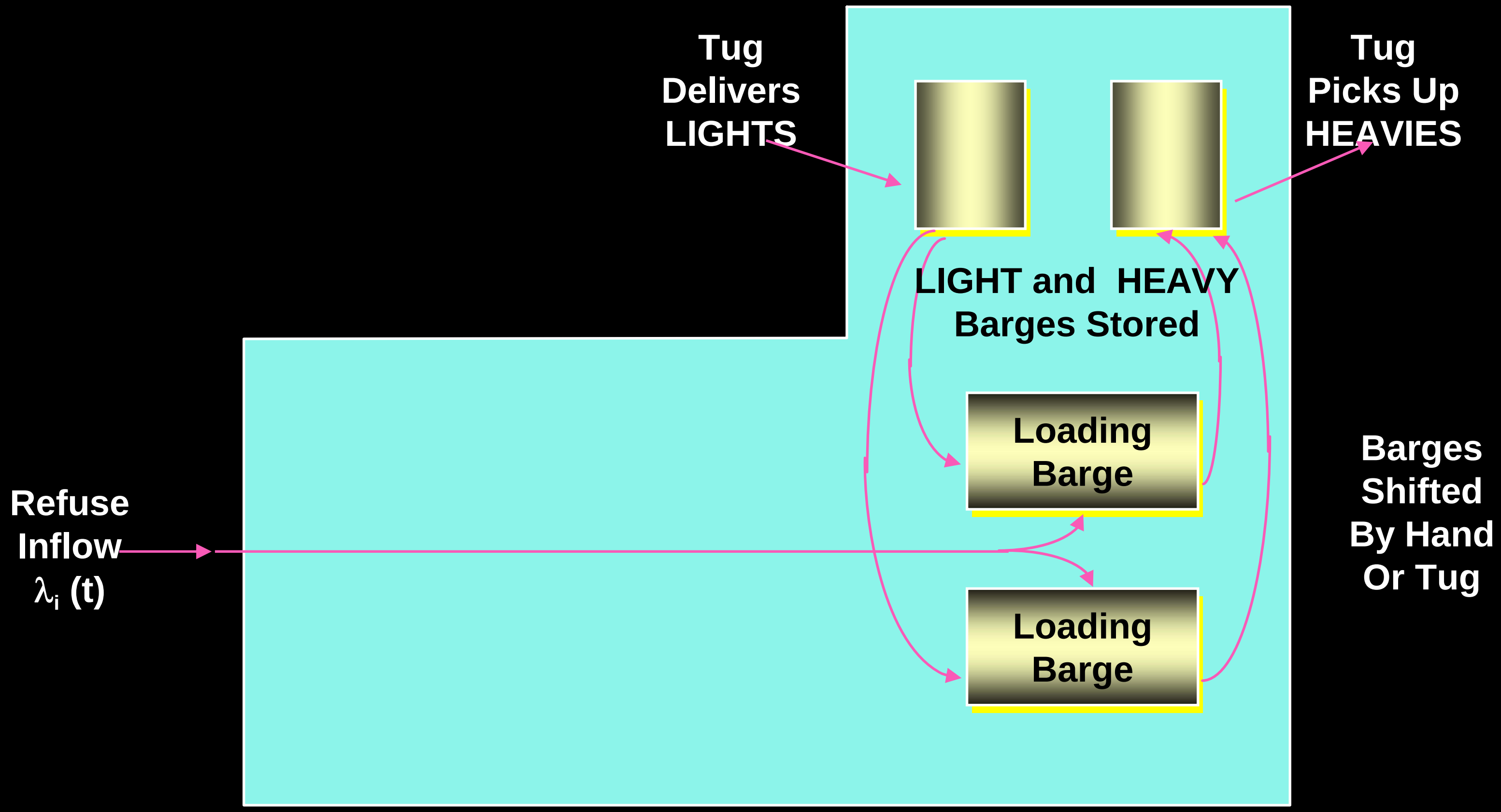
Photo courtesy of Eddie Codel.
http://www.flickr.com/photos/ekai/15899569/

Marine Transfer Station

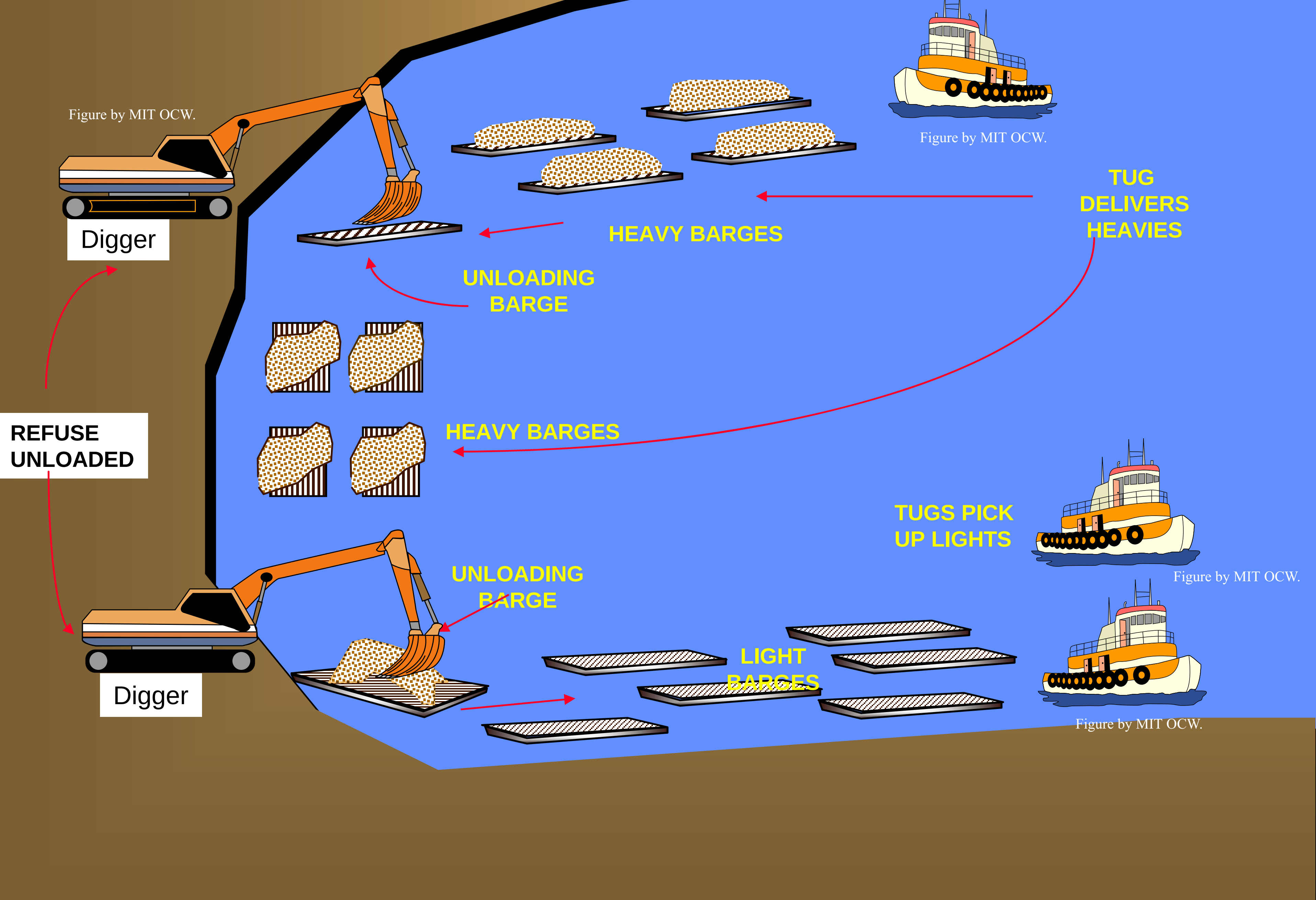
Courtesy of Dattner Architects. Used with permission.

http://www.dattner.com/html/civic1a.html

NYC Marine Transfer Station



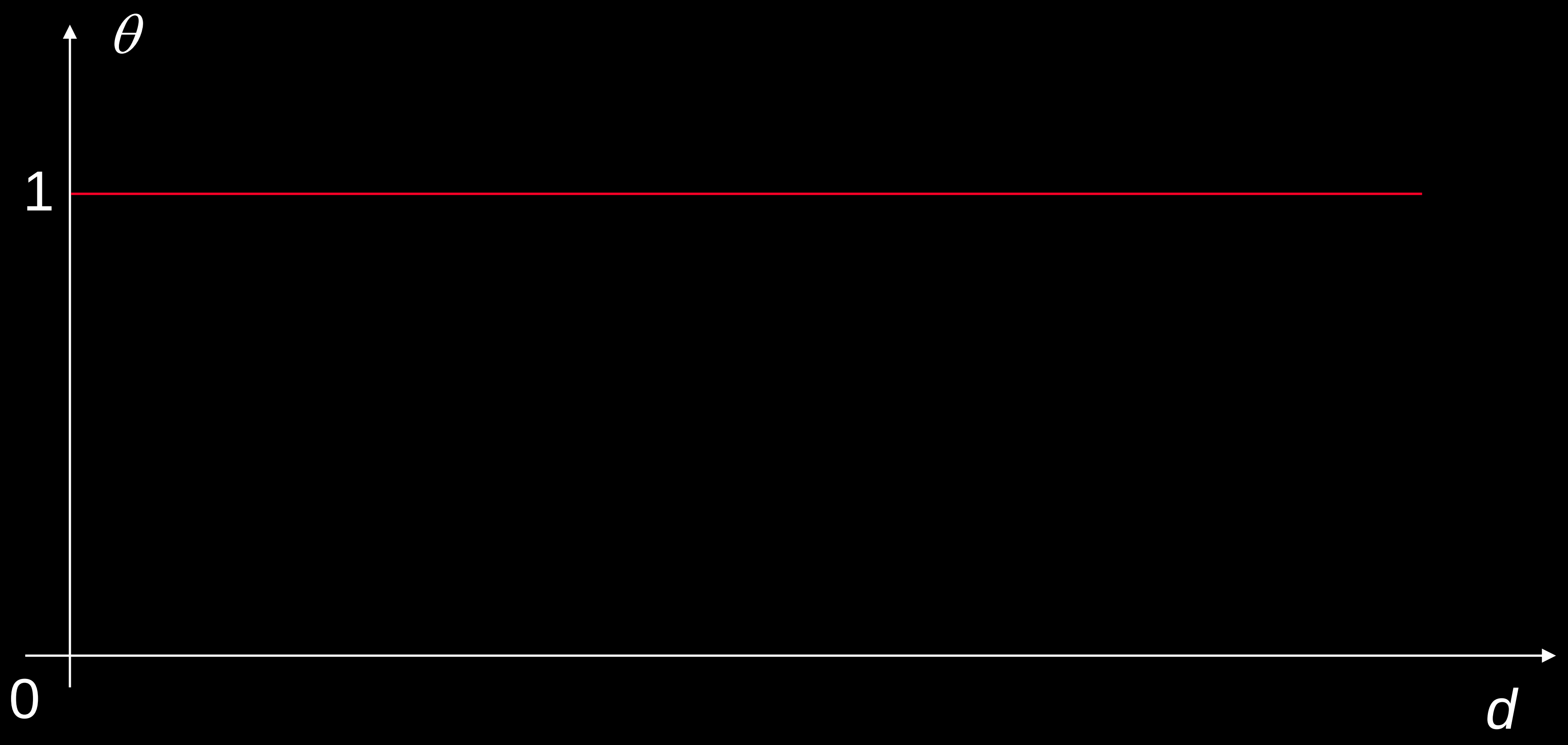
Fresh Kills Landfill

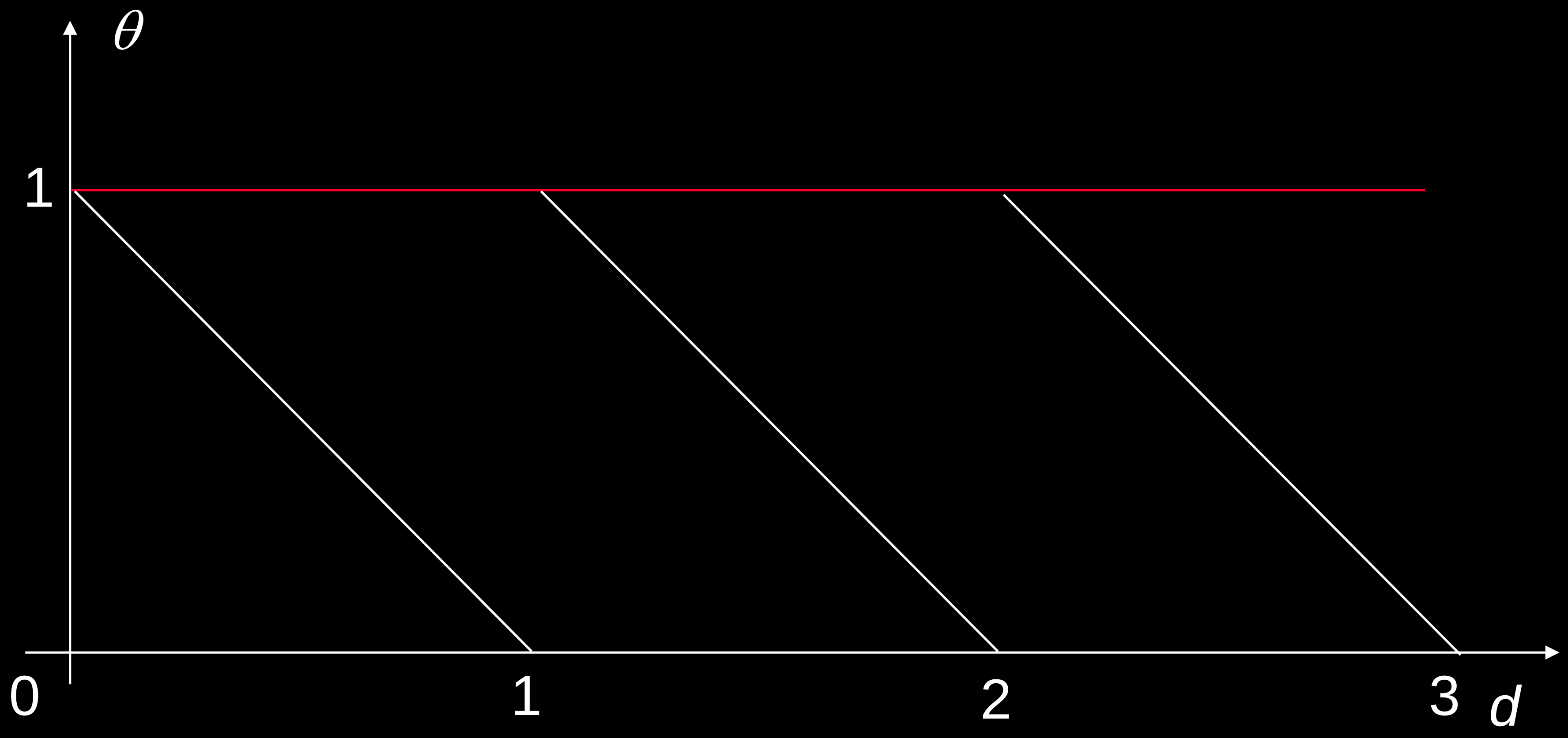


1. The R.V.'s

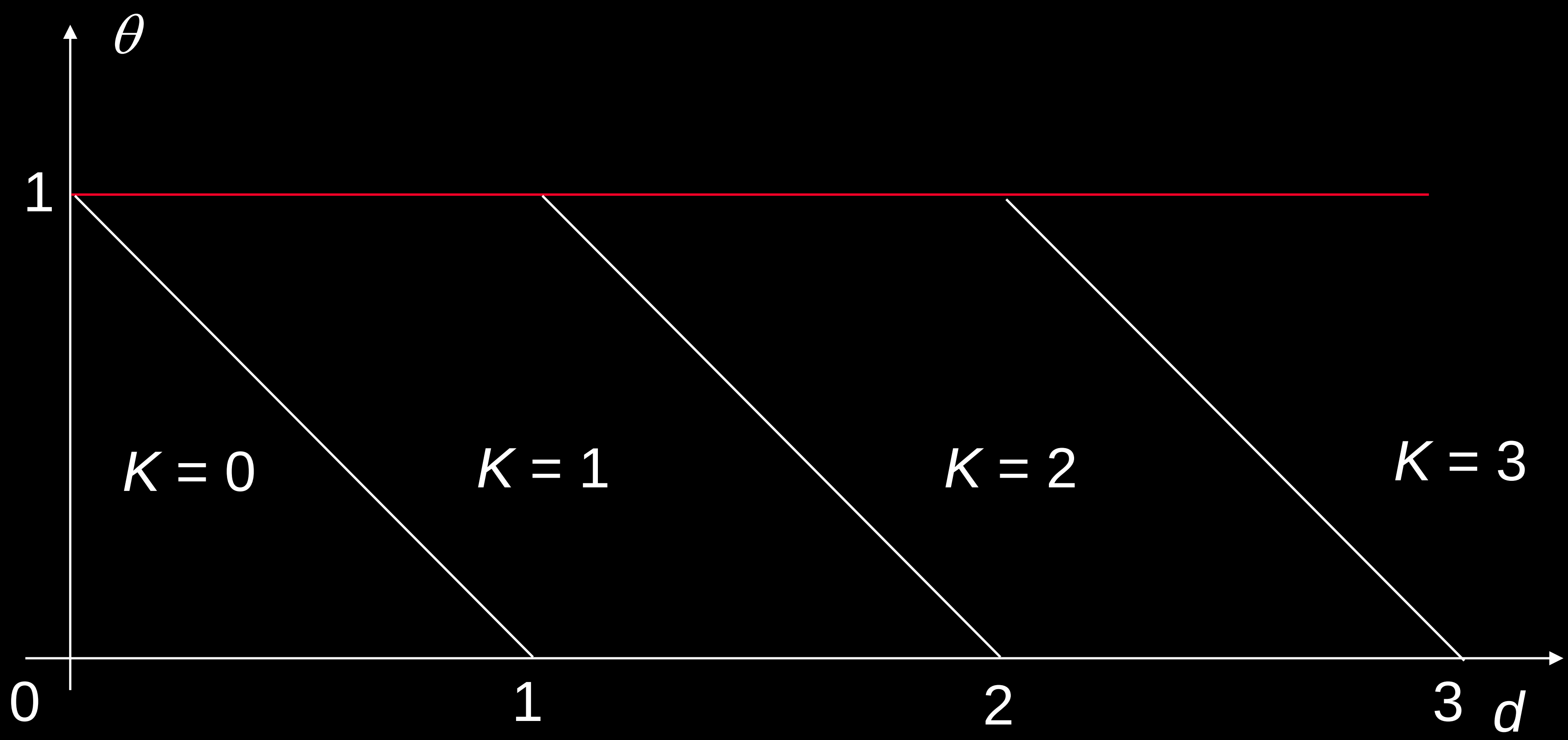
- D = barge loads of garbage produced on a random day (continuous r.v.)
- Θ = fraction of barge that is filled at beginning of day ($0 < \Theta < 1$)
- K = total number of completely filled barges produced by a facility on a random day (K integer)
- $K = \lfloor D + \Theta \rfloor$ = integer part of $D + \Theta$

2. The Sample Space





θ
1
0
1
2
3
d


θ
1
K = 0
K = 1
K = 2
K = 3
0
1
2
3
d

3. Joint Probability Distribution

a) D and Θ are independent.

b) Θ is uniformly distributed over [0, 1]

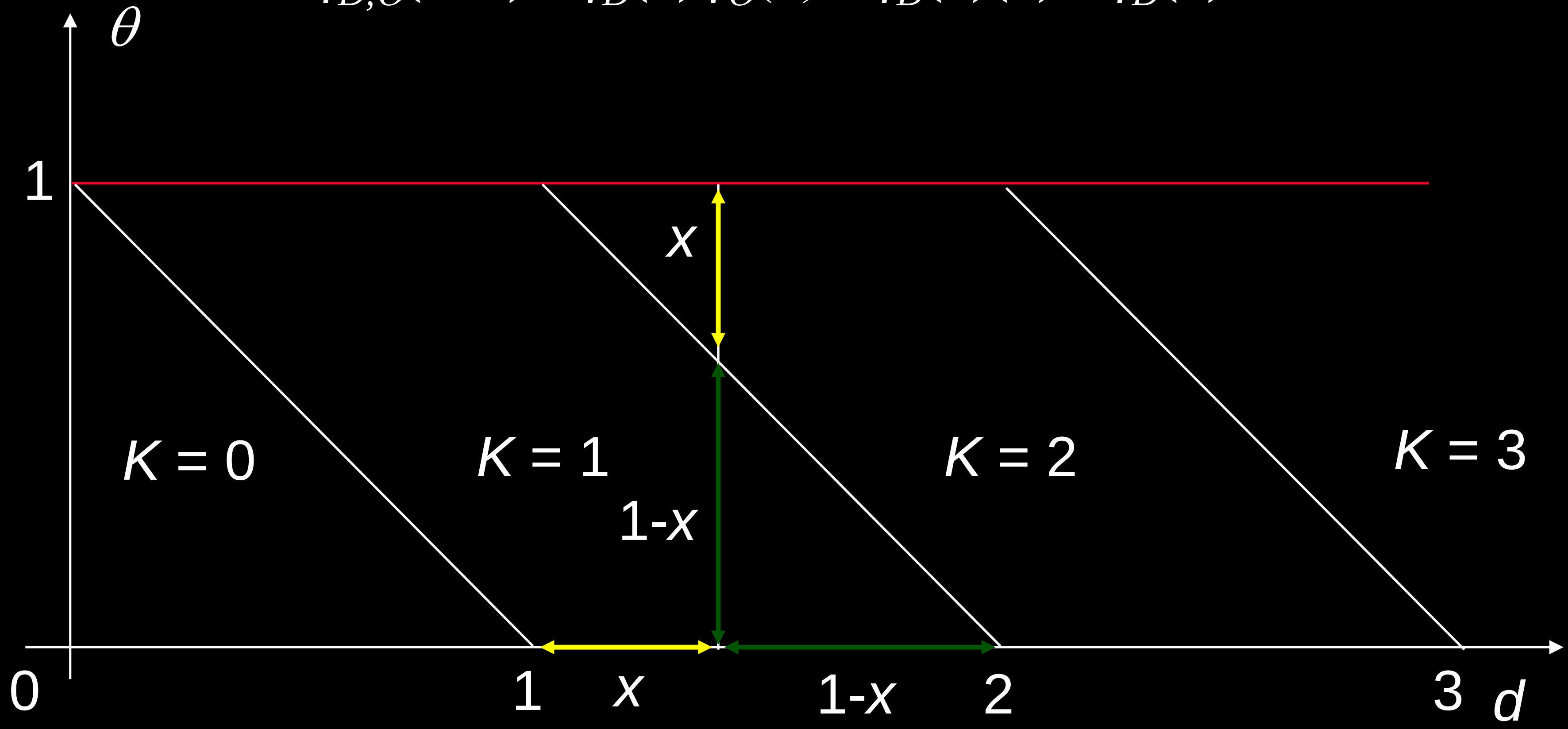
$$f_{D,\Theta}(d, \theta) = f_D(d)\, f_\Theta(\theta) = f_D(d)(1) = f_D(d),\ d > 0,\ 0 < \theta < 1$$

3. Joint Probability Distribution

a) D and Θ are independent.

b) Θ is uniformly distributed over [0, 1]

$$f_{D,\Theta}(d, \theta) = f_D(d)\, f_\Theta(\theta) = f_D(d)(1) = f_D(d),\ d > 0,\ 0 < \theta < 1$$



4. Working in the Joint Sample Space

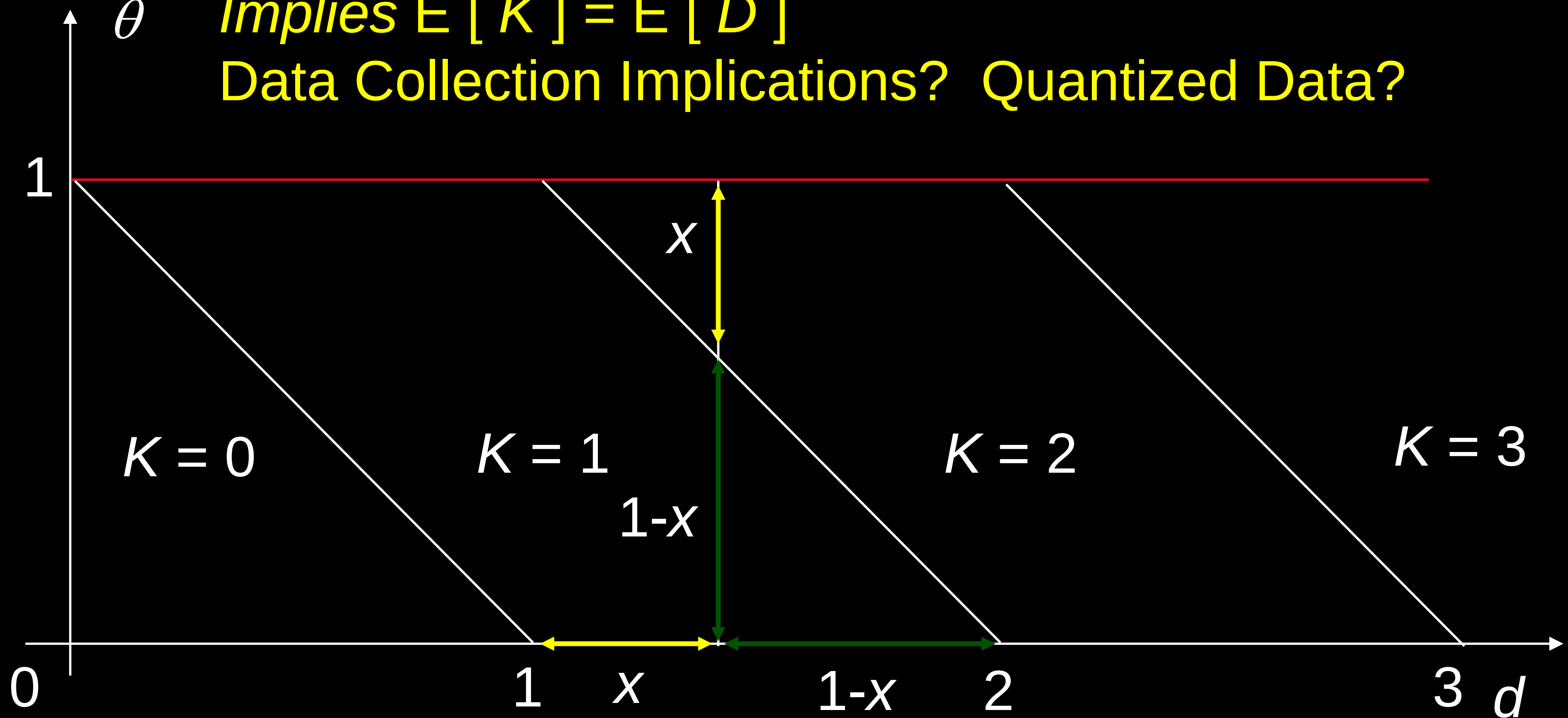
Look at $E[K \mid D = d]$

Let $d = i + x \quad 0 < x < 1$

$E[K \mid D = i + x] = i(1 - x) + (i + 1)x = i + x = d$

Implies $E[K] = E[D]$

Data Collection Implications? Quantized Data?



What Have We Learned Today?

4 Steps: Functions of R.V.s

1. Define the Random Variables
2. Identify the joint sample space
3. Determine the probability law over the sample space
4. Carefully work in the sample space to answer any question of interest

4a. Derive the CDF of the R.V. of interest, working in the original sample space whose probability law you know

4b Take the derivative to obtain the desired PDF