Conditional PDF, given an event

$$p_X(x) = \mathbf{P}(X = x)$$

$$f_X(x) \cdot \delta \approx \mathbf{P}(x \le X \le x + \delta)$$

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A)$$

$$f_{X|A}(x) \cdot \delta \approx \mathbf{P}(x \le X \le x + \delta \mid A)$$

Conditional PDF, given an event

$$p_X(x) = \mathbf{P}(X = x)$$

$$f_X(x) \cdot \delta \approx \mathbf{P}(x \leq X \leq x + \delta)$$

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A)$$

$$f_{X|A}(x) \cdot \delta \approx \mathbf{P}(x \leq X \leq x + \delta \mid A)$$

$$\mathbf{P}(X \in B) = \sum_{x \in B} p_X(x)$$

$$\mathbf{P}(X \in B) = \int_B f_X(x)\, dx$$

$$\mathbf{P}(X \in B \mid A) = \sum_{x \in B} p_{X|A}(x)$$

$$\mathbf{P}(X \in B \mid A) = \int_B f_{X|A}(x)\, dx$$

Conditional PDF, given an event

$$p_X(x) = \mathbf{P}(X = x)$$

$$f_X(x) \cdot \delta \approx \mathbf{P}(x \le X \le x + \delta)$$

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A)$$

$$f_{X|A}(x) \cdot \delta \approx \mathbf{P}(x \le X \le x + \delta \mid A)$$

$$\mathbf{P}(X \in B) = \sum_{x \in B} p_X(x)$$

$$\mathbf{P}(X \in B) = \int_B f_X(x)\, dx$$

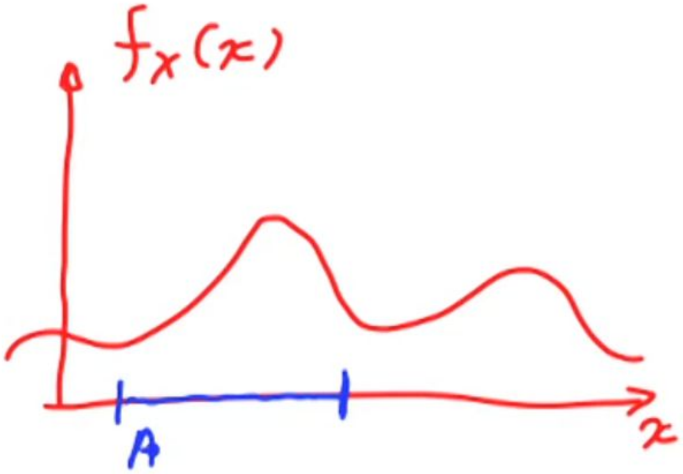
$$\mathbf{P}(X \in B \mid A) = \sum_{x \in B} p_{X|A}(x)$$

$$\mathbf{P}(X \in B \mid A) = \int_B f_{X|A}(x)\, dx$$

$$\sum_x p_{X|A}(x) = 1$$

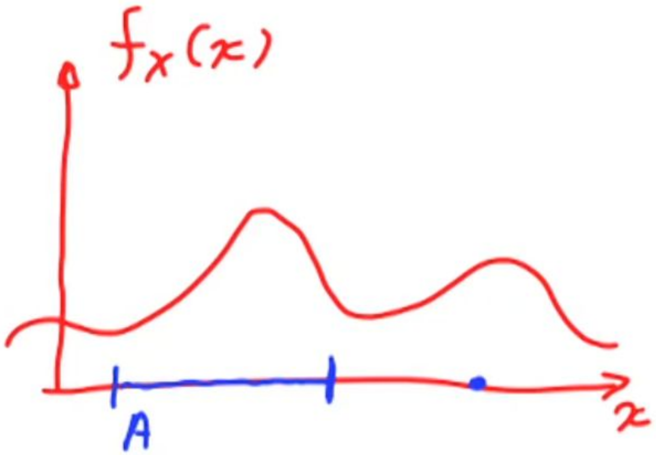
$$\int f_{X|A}(x)\, dx = 1$$

Conditional PDF of X, given that $X \in A$



Conditional PDF of X, given that $X \in A$

$$\mathbf{P}(x \le X \le x + \delta \mid X \in A) \approx f_{X|X \in A}(x) \cdot \delta$$

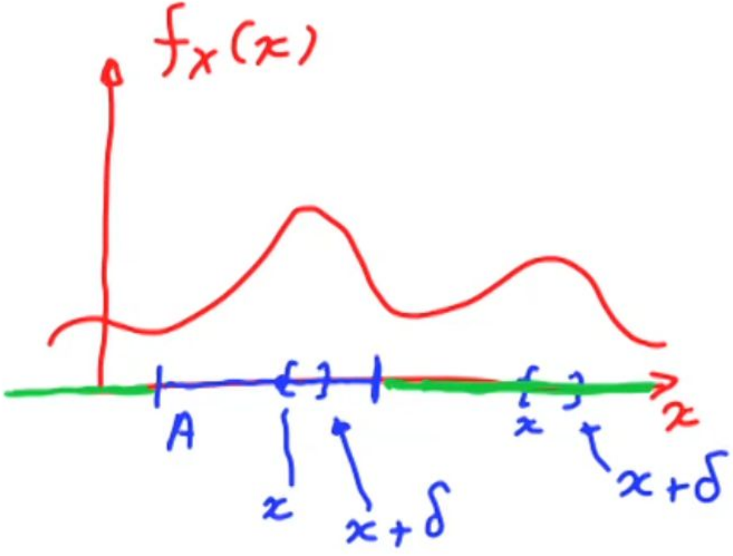


Conditional PDF of X, given that $X \in A$

$$\mathbf{P}(x \le X \le x + \delta \mid X \in A) \approx f_{X|X \in A}(x) \cdot \cancel{\delta}$$

$$= \frac{\mathbf{P}(x \le X \le x + \delta,\ X \in A)}{\mathbf{P}(A)}$$

$$= \frac{\mathbf{P}(x \le X \le x + \delta)}{\mathbf{P}(A)} \approx \frac{f_X(x)\,\cancel{\delta}}{\mathbf{P}(A)}$$

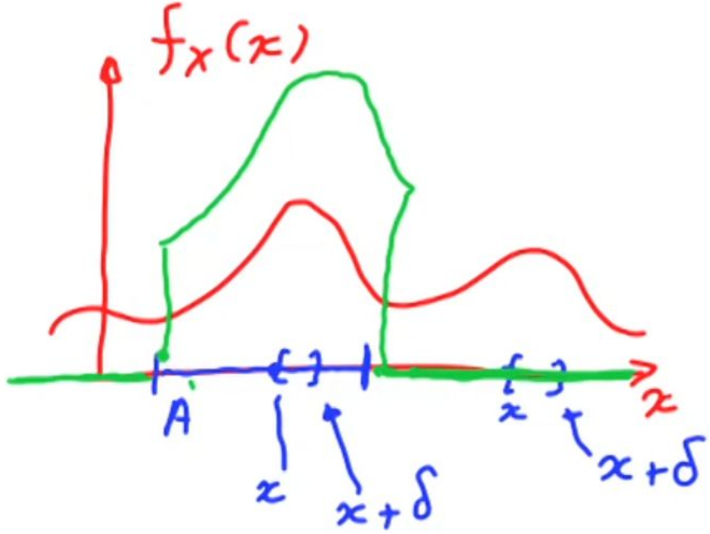


Conditional PDF of X, given that $X \in A$

$$\mathbf{P}(x \le X \le x + \delta \mid X \in A) \approx f_{X|X\in A}(x) \cdot \cancel{\delta}$$

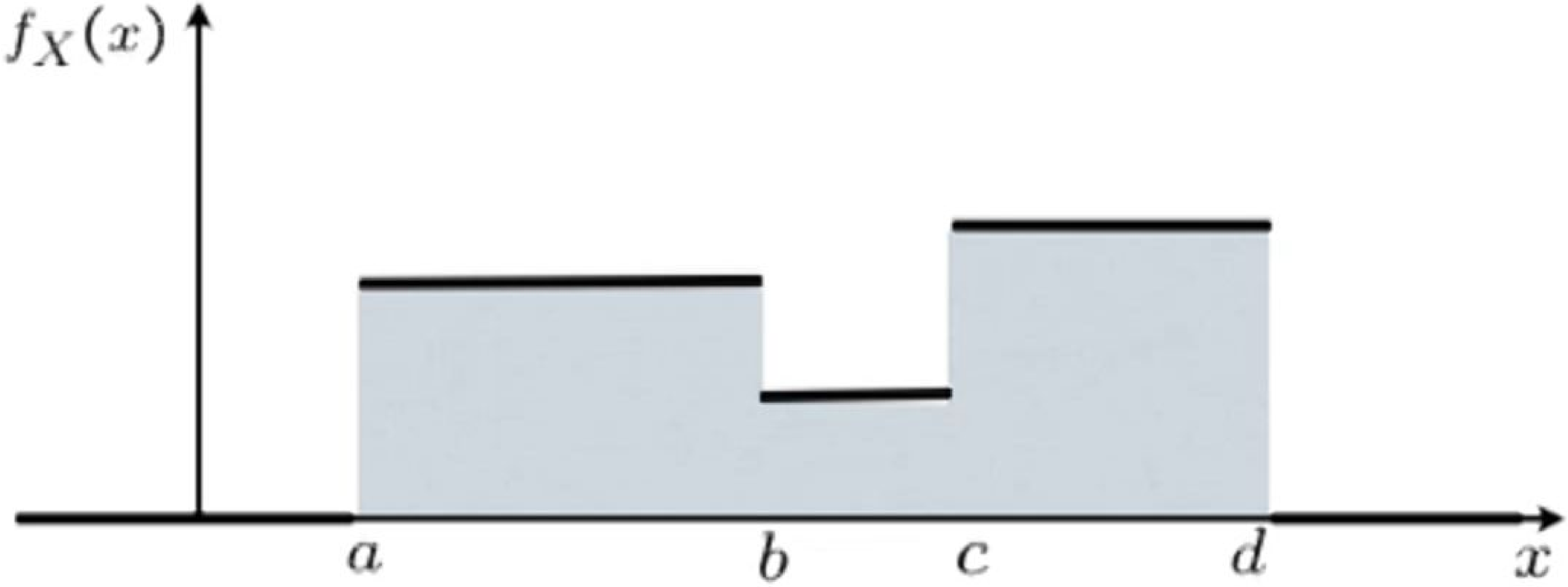
$$= \frac{\mathbf{P}(x \le X \le x+\delta,\ X \in A)}{\mathbf{P}(A)}$$

$$= \frac{\mathbf{P}(x \le X \le x+\delta)}{\mathbf{P}(A)} \approx \frac{f_X(x)\,\cancel{\delta}}{\mathbf{P}(A)}$$



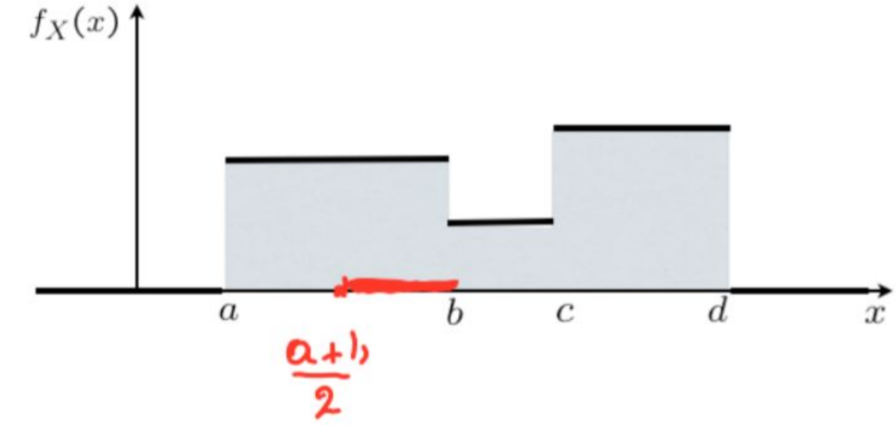
$$f_{X|X\in A}(x) = \begin{cases} 0, & \text{if } x \notin A \\ \dfrac{f_X(x)}{\mathbf{P}(A)}, & \text{if } x \in A \end{cases}$$

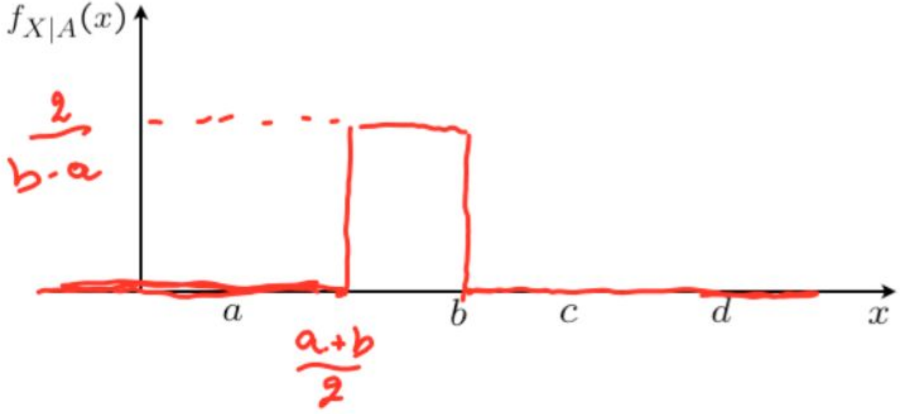
Example



Example

$$A: \quad \frac{a+b}{2} \le X \le b$$





$$\mathbf{E}[X \mid A] = \frac{1}{2} \cdot \frac{a+b}{2} + \frac{1}{2} b$$

$$= \frac{1}{4} a + \frac{3}{4} b$$

$$\mathbf{E}[X^2 \mid A] = \int_{\frac{a+b}{2}}^{b} \frac{2}{b-a} \cdot x^2 \, dx$$

Memorylessness of the exponential PDF

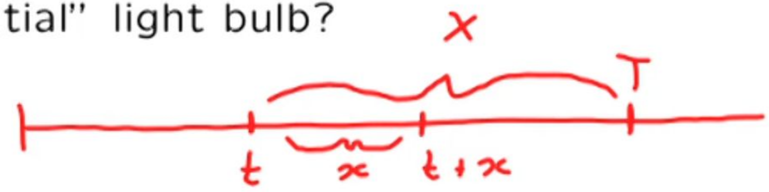
- Do you prefer a used or a new "exponential" light bulb?
- Bulb lifetime T: exponential(λ)

  $\mathbf{P}(T > x) = e^{-\lambda x}$, for $x \geq 0$

  - we are told that $T > t$
  - r.v. X: remaining lifetime

Memorylessness of the exponential PDF

- Do you prefer a used or a new "exponential" light bulb?



- Bulb lifetime T: exponential(λ)

  $\mathbf{P}(T > x) = e^{-\lambda x}$, for $x \geq 0$

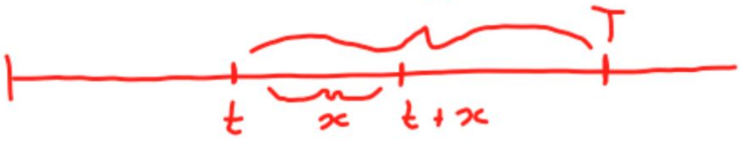
  – we are told that $T > t$

  – r.v. X: remaining lifetime $= T - t$

  $\mathbf{P}(X > x \,|\, T > t) =$

Memorylessness of the exponential PDF

- Do you prefer a used or a new "exponential" light bulb? **Probabilistically identical!**



- Bulb lifetime T: exponential(λ)

  $\mathbf{P}(T > x) = e^{-\lambda x}$, for $x \geq 0$

  – we are told that $T > t$

  – r.v. X: remaining lifetime $= T - t$

  $\mathbf{P}(X > x \,|\, T > t) = e^{-\lambda x}$, for $x \geq 0$

$$= \frac{\mathbf{P}(T - t > x,\ T > t)}{\mathbf{P}(T > t)} = \frac{\mathbf{P}(T > t + x,\ T > t)}{\mathbf{P}(T > t)} = \frac{\mathbf{P}(T > t + x)}{\mathbf{P}(T > t)}$$

$$= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x}$$

Memorylessness of the exponential PDF

$f_T(x) = \lambda e^{-\lambda x}$, for $x \geq 0$

$\mathbf{P}(0 \leq T \leq \delta)$

$\mathbf{P}(t \leq T \leq t + \delta \mid T > t)$

similar to an independent coin flip,
every δ time steps,
with $\mathbf{P}(\text{success}) \approx \lambda\delta$

Memorylessness of the exponential PDF

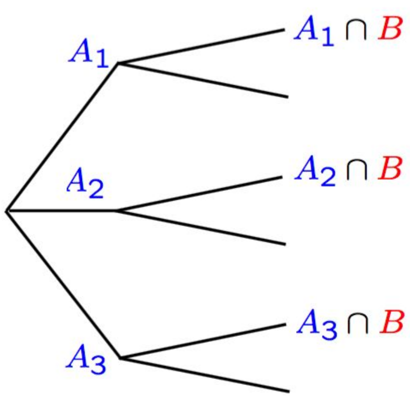
$$f_T(x) = \lambda e^{-\lambda x}, \quad \text{for } x \geq 0$$

$$\mathbf{P}(0 \leq T \leq \delta) \approx f_T(0) \cdot \delta = \lambda\delta$$

$$\mathbf{P}(t \leq T \leq t + \delta \mid T > t) = \quad \approx \lambda\delta$$

similar to an independent coin flip,
every δ time steps,
with $\mathbf{P}(\text{success}) \approx \lambda\delta$

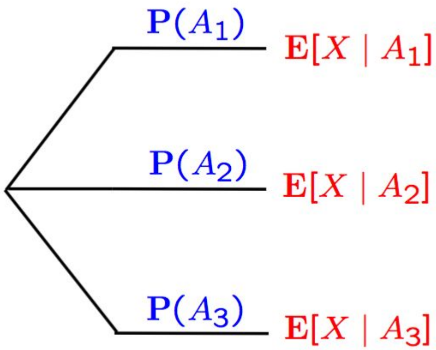
Total probability and expectation theorems



$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)$$

$$p_X(x) = \mathbf{P}(A_1)\, p_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)\, p_{X|A_n}(x)$$

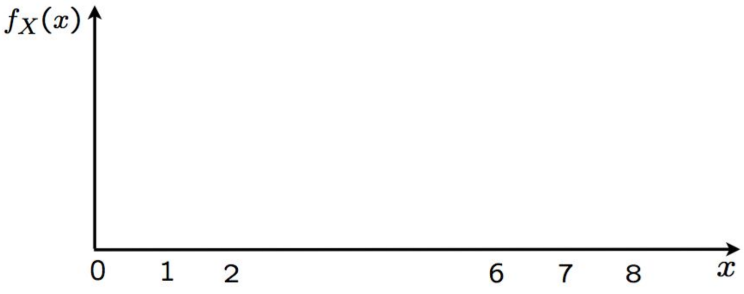
$$\boxed{f_X(x) = \mathbf{P}(A_1) f_{X|A_1}(x) + \cdots + \mathbf{P}(A_n) f_{X|A_n}(x)}$$



$$\boxed{\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X \mid A_n]}$$

Example

- Bill goes to the supermarket shortly, with probability 1/3, at a time uniformly distributed between 0 and 2 hours from now; or with probability 2/3, later in the day at a time uniformly distributed between 6 and 8 hours from now



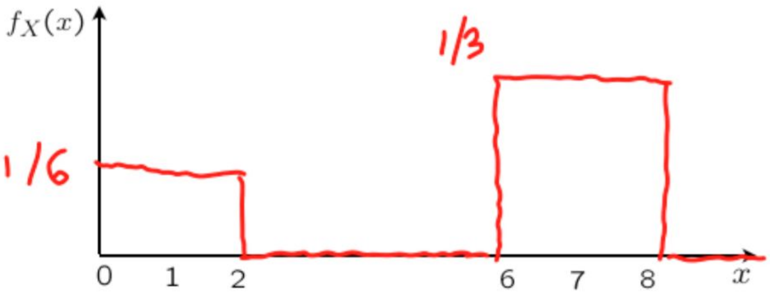
$$f_X(x) = \mathbf{P}(A_1) f_{X|A_1}(x) + \cdots + \mathbf{P}(A_n) f_{X|A_n}(x)$$

$$\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X \mid A_n]$$

Example

- Bill goes to the supermarket shortly, with probability 1/3, at a time uniformly distributed between 0 and 2 hours from now; or with probability 2/3, later in the day at a time uniformly distributed between 6 and 8 hours from now

$P(A_1) = \frac{1}{3}$ $\quad f_{X|A_1} \sim \text{unif}[0,2]$ $\quad P(A_2) = \frac{2}{3}$ $\quad f_{X|A_2} \sim U[6,8]$



$$f_X(x) = \mathbf{P}(A_1) f_{X|A_1}(x) + \cdots + \mathbf{P}(A_n) f_{X|A_n}(x)$$

- $\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X \mid A_n]$

$$\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 7$$

Jointly continuous r.v.'s and joint PDFs

$p_X(x)$	$f_X(x)$
$p_{X,Y}(x,y)$	$f_{X,Y}(x,y)$

$$p_{X,Y}(x,y) = \mathbf{P}(X = x \text{ and } Y = y) \geq 0$$

$$f_{X,Y}(x,y) \geq 0$$

$$\mathbf{P}\Big((X,Y) \in B\Big) = \sum_{(x,y)\in B}\sum p_{X,Y}(x,y)$$

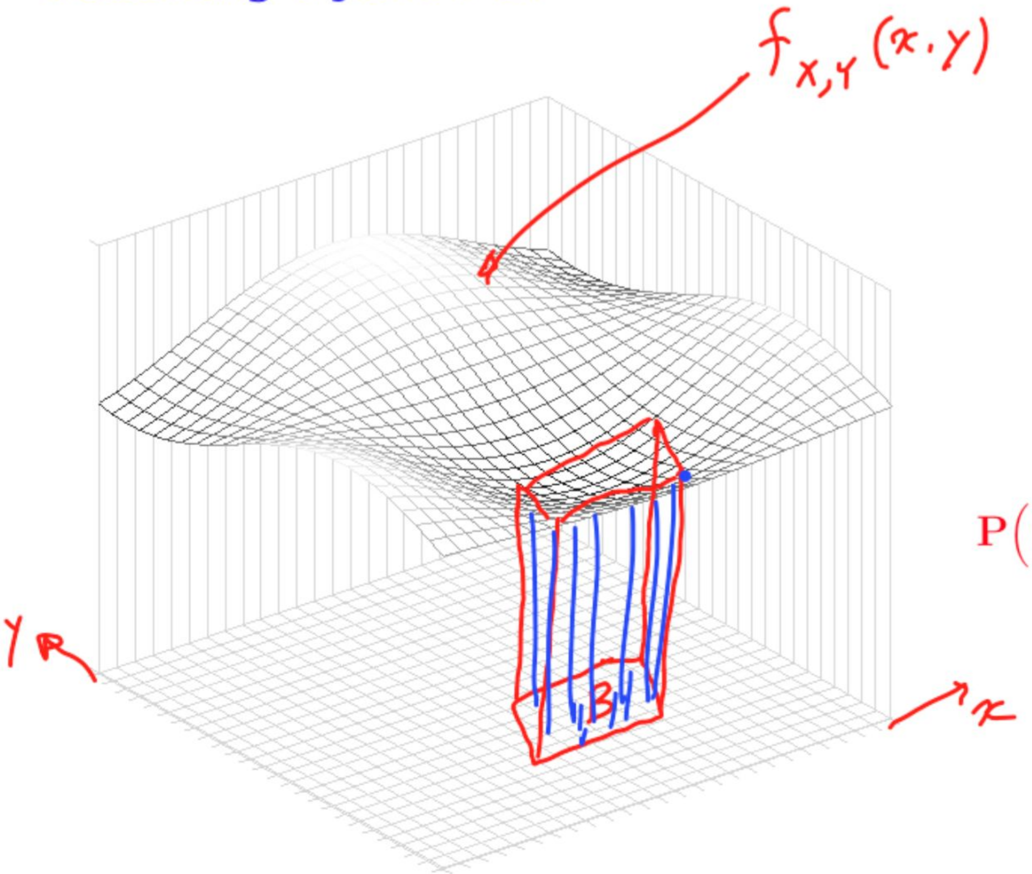
$$\mathbf{P}\Big((X,Y) \in B\Big) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

$$\sum_x \sum_y p_{X,Y}(x,y) = 1$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

Definition: Two random variables are **jointly continuous** if they can be described by a joint PDF

Visualizing a joint PDF



$$\mathbf{P}\Big((X,Y) \in B\Big) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

On joint PDFs

$$\mathbf{P}\big((X,Y) \in B\big) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

$$\mathbf{P}(a \le X \le b,\, c \le Y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\,dx\,dy$$

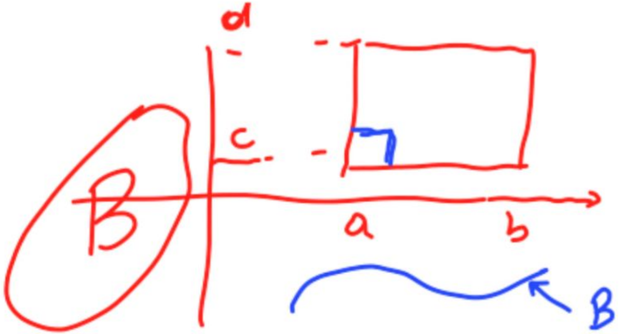
$$\boxed{\mathbf{P}(a \le X \le a+\delta,\, c \le Y \le c+\delta) \approx f_{X,Y}(a,c) \cdot \delta^2}$$

$f_{X,Y}(x,y)$: probability per unit area

area$(B) = 0 \quad \Rightarrow \quad \mathbf{P}\big((X,Y) \in B\big) = 0$

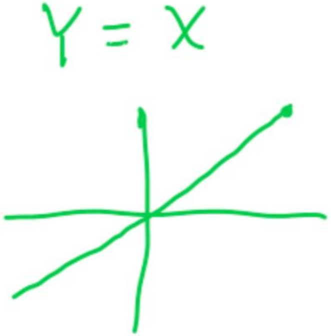
On joint PDFs

$$\mathbf{P}\Big((X,Y)\in B\Big) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy$$



$$\mathbf{P}(a\le X\le b,\ c\le Y\le d) = \int_c^d\int_a^b f_{X,Y}(x,y)\,dx\,dy$$

$$\boxed{\mathbf{P}(a\le X\le a+\delta,\ c\le Y\le c+\delta)\approx f_{X,Y}(a,c)\cdot\delta^2}$$



$f_{X,Y}(x,y)$: probability per unit area

area$(B)=0 \quad\Rightarrow\quad \mathbf{P}\Big((X,Y)\in B\Big)=0$

From the joint to the marginals

$$p_X(x) = \sum_y p_{X,Y}(x,y) \qquad f_X(x) = \int f_{X,Y}(x,y)\,dy$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y) \qquad f_Y(y) = \int f_{X,Y}(x,y)\,dx$$

Joint, Marginal and Conditional Densities

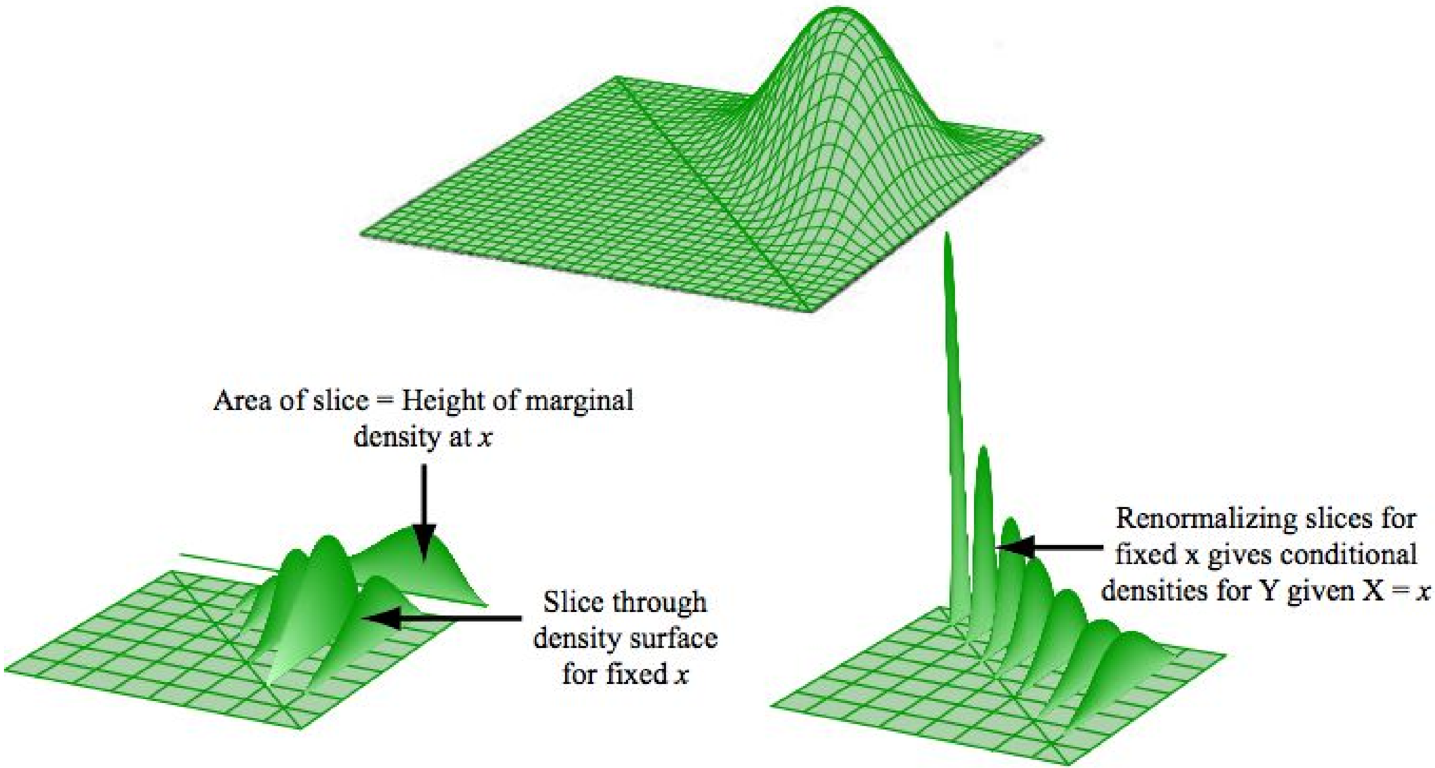


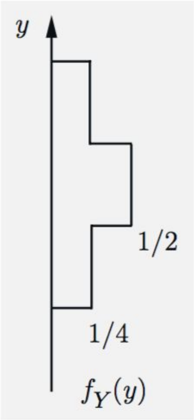
Image adapted from
Probability, by J. Pittman, 1999.

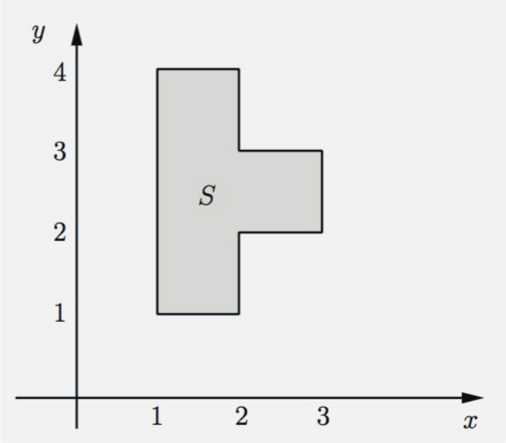
Uniform joint PDF on a set S

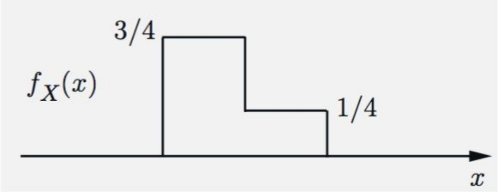
$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\text{area of } S}, & \text{if } (x,y) \in S, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_X(x) = \int f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int f_{X,Y}(x,y)\,dx$$



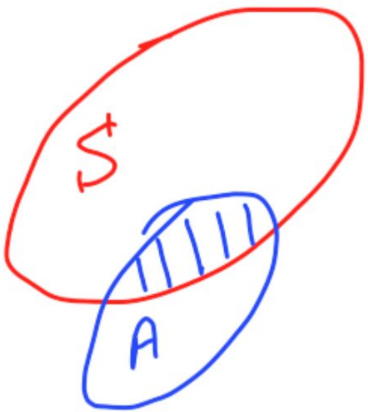




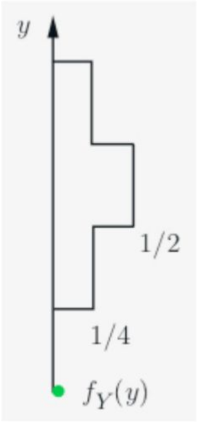
Uniform joint PDF on a set S

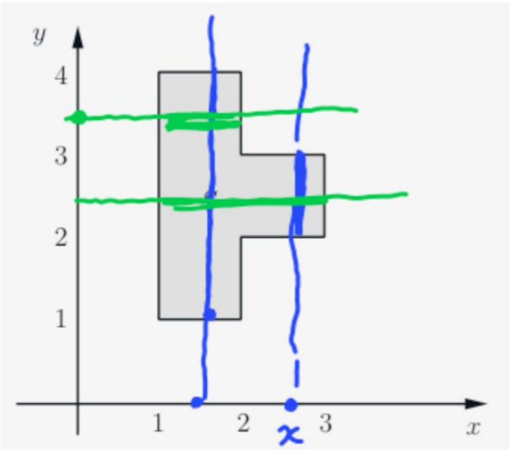
$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\text{area of } S}, & \text{if } (x,y) \in S, \\ 0, & \text{otherwise.} \end{cases}$$

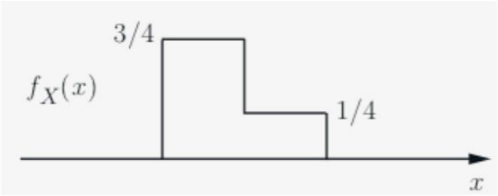
$f_{x,y} = \frac{1}{4}$



$P(A) = \dfrac{\text{area}(A \cap S)}{\text{area}(S)}$







Conditional PDFs, given another r.v.

$$p_{X,Y}(x,y) \qquad f_{X,Y}(x,y)$$
$$p_{X|A}(x) \qquad f_{X|A}(x)$$
$$p_{X|Y}(x \mid y) \qquad f_{X|Y}(x \mid y)$$

$$p_{X|Y}(x\,|\,y) = \mathbf{P}(X = x\,|\,Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}, \quad \text{if } p_Y(y) > 0$$

Definition: $$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{if } f_Y(y) > 0$$

$$\mathbf{P}(x \le X \le x + \delta\,|\,A) \approx f_{X|A}(x) \cdot \delta, \qquad \text{where } \mathbf{P}(A) > 0$$

$$\mathbf{P}(x \le X \le x + \delta \mid y \le Y \le y + \epsilon)$$

Definition: $$\mathbf{P}(X \in A \mid Y = y) = \int_A f_{X\,|\,Y}(x \mid y)\,dx$$

Comments on conditional PDFs

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

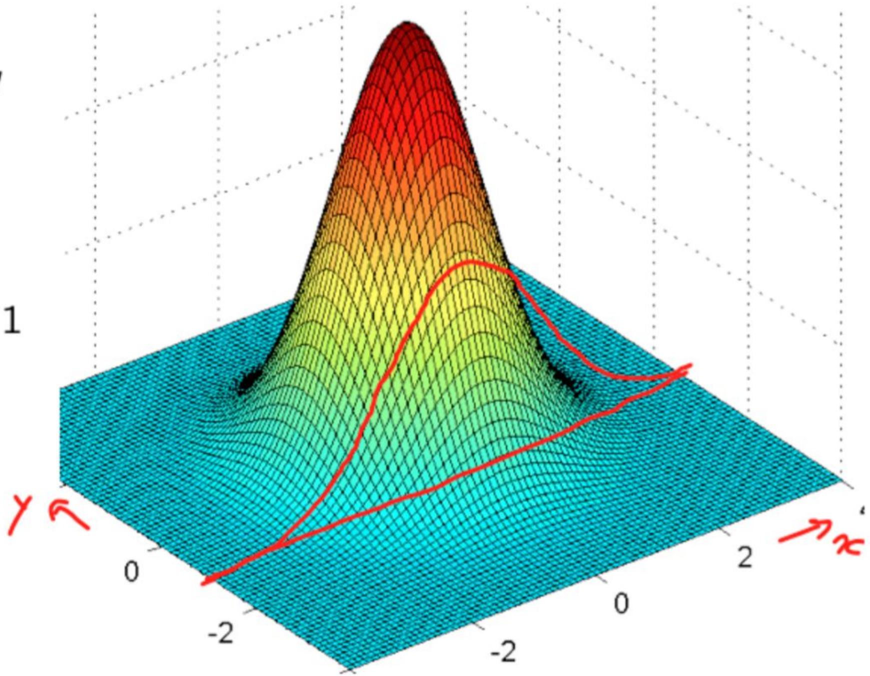
- $f_{X|Y}(x \mid y) \geq 0$

- Think of value of Y as fixed at some y
  shape of $f_{X|Y}(\cdot \mid y)$: slice of the joint

- $\displaystyle\int_{-\infty}^{\infty} f_{X|Y}(x \mid y)\,dx = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx}{f_Y(y)} = 1$

- Multiplication rule:

$$f_{X,Y}(x,y) = f_Y(y) \cdot f_{X|Y}(x \mid y)$$
$$= f_X(x) \cdot f_{Y|X}(y \mid x)$$



Independence

$p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$, for all x, y

$\boxed{f_{X,Y}(x,y) = f_X(x)\,f_Y(y), \quad \text{for all } x \text{ and } y}$

$f_{X,Y}(x,y) = f_{X|Y}(x\,|\,y)\,f_Y(y)$

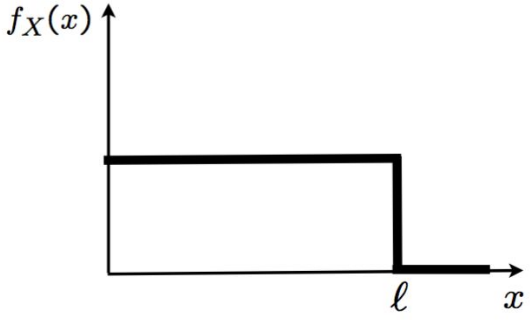
- equivalent to: $f_{X|Y}(x\,|\,y) = f_X(x)$, for all y with $f_Y(y) > 0$ and all x

If X, Y are **independent:** $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$

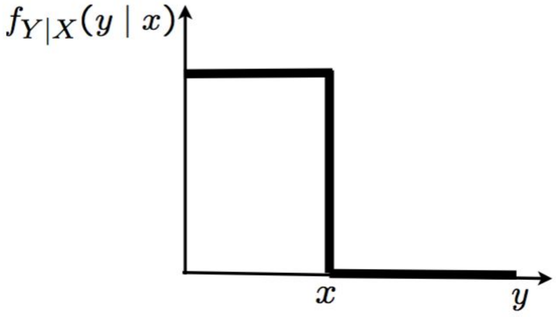
$g(X)$ and $h(Y)$ are also independent: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

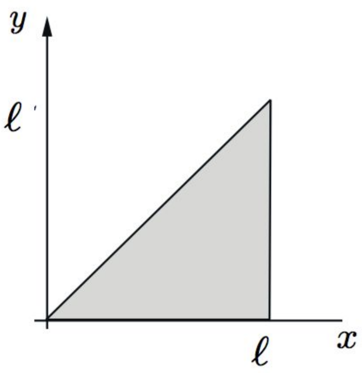
Stick-breaking example



- Break a stick of length ℓ twice
  - first break at X: uniform in $[0, \ell]$
  - second break at Y: uniform in $[0, X]$

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y \mid x) =$$



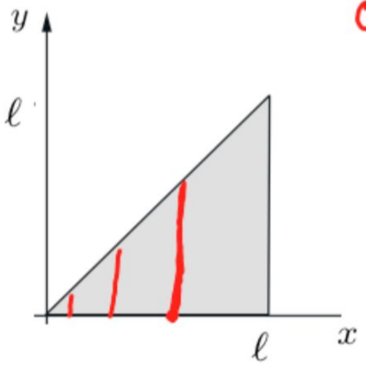


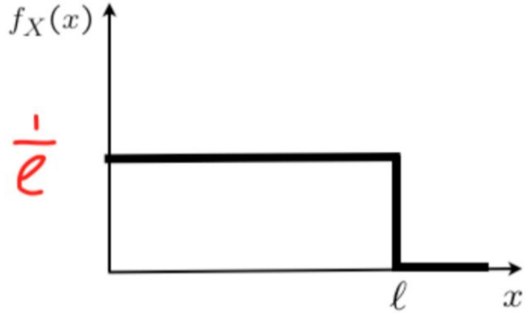
Stick-breaking example

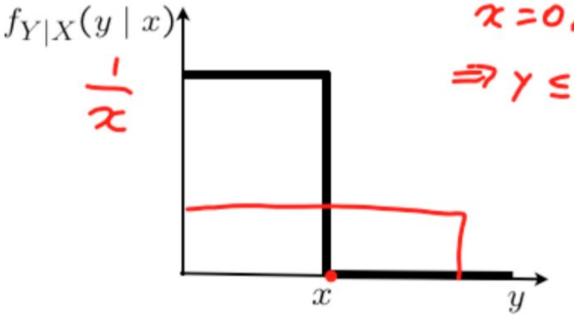
- Break a stick of length ℓ twice
  - first break at X: uniform in $[0, \ell]$
  - second break at Y: uniform in $[0, X]$

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y \mid x) = \frac{1}{\ell x}$$

$$0 \le y \le x \le \ell$$



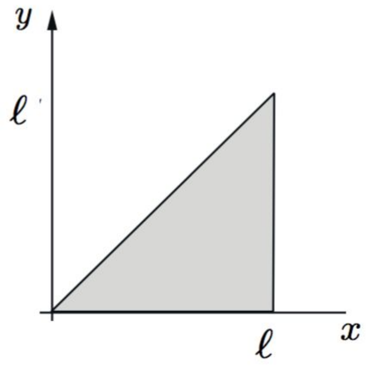




$x = 0.5$

$\Rightarrow y \le 0.5$

Stick-breaking example



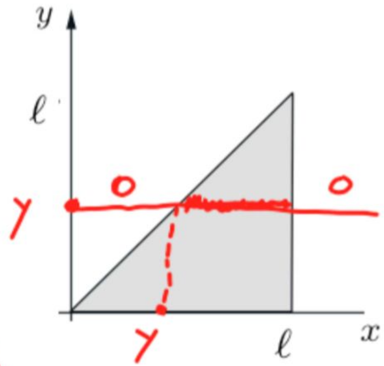
$$f_{X,Y}(x,y) = \frac{1}{\ell x}, \qquad 0 \le y \le x \le \ell$$

$$f_Y(y) \; =$$

$$\mathbf{E}[Y] =$$

- Using total expectation theorem:

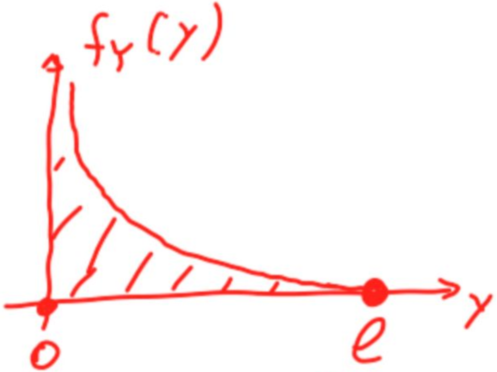
Stick-breaking example



$$f_{X,Y}(x,y) = \frac{1}{\ell x}, \qquad 0 \le y \le x \le \ell$$

$$f_Y(y) = \int f_{X,Y}(x,y)\,dx = \int_y^{\ell} \frac{1}{\ell x}\,dx = \frac{1}{\ell}\log\left(\frac{\ell}{y}\right)$$

$$\mathbf{E}[Y] = \int_0^{\ell} y \frac{1}{\ell}\log\left(\frac{\ell}{y}\right) dy$$



- Using total expectation theorem:

$$E[Y] = \int_0^{\ell} \frac{1}{\ell} E[Y \mid X = x]\,dx = \int_0^{\ell} \underbrace{\frac{1}{\ell}}_{f_X(x)} \cdot \frac{x}{2}\,dx = \frac{1}{2}E[X] = \frac{1}{2}\cdot\frac{\ell}{2} = \frac{\ell}{4}.$$