Further Topics on Random Variables: Convolution, Conditional Expectation and Variance

Berlin Chen
Department of Computer Science & Information Engineering
National Taiwan Normal University

Reference:

- D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability* , Sections 4.2-3

Sums of Independent Random Variables (1/2)

- **Recall:** If X and Y are independent random variables, the distribution (PMF or PDF) of $W = X + Y$ can be obtained by computing and inverting the **transform**

$$M_W(s) = M_X(s) M_Y(s)$$

- We also can use the **convolution** method to obtain the distribution of $W = X + Y$
  - If X and Y are independent **discrete random variables** with integer values

$$\begin{aligned} p_W(w) &= \mathbf{P}(X + Y = w) = \sum_{\{(x,y) \mid x+y=w\}} \mathbf{P}(X = x, Y = y) \\ &= \sum_x \mathbf{P}(X = x, Y = w - x) = \sum_x \mathbf{P}(X = x) \mathbf{P}(Y = w - x) \\ &= \boxed{\sum_x p_X(x) p_Y(w - x)} \quad \left(\text{also equivalent to} \sum_y p_X(w - y) p_Y(y) \right) \end{aligned}$$

Convolution of PMFs of X and Y

Sums of Independent Random Variables (2/2)

– If X and Y are independent **continuous random variables**, the PDF $f_W(w)$ of $W = X + Y$ can be obtained by

$$f_W(w) = \boxed{\int_{-\infty}^{\infty} f_X(x) f_Y(w-x)\,dx}$$

Convolution of PMFs of X and Y

$$\left(\text{also equivalent to } \int_{-\infty}^{\infty} f_X(w-x) f_Y(y)\,dy\right)$$

Note that

$$\mathbf{P}(W \le w \mid X = x) = \mathbf{P}(X + Y \le w \mid X = x)$$
$$= \mathbf{P}(x + Y \le w)$$
$$= \mathbf{P}(Y \le w - x)$$

independence assumption

$$\Rightarrow F_{W|X}(w|x) = F_Y(w-x)$$

Differentiate the CDFs of both sides with respect to w

$$\Rightarrow f_{W|X}(w|x) = f_Y(w-x)$$

Applying the multiplication (chain) rule, we have

$$f_{W,X}(w,x) = f_X(x) f_{W|X}(w|x)$$
$$= f_X(x) f_Y(w-x)$$

Finally, by marginalization, we can have

$$f_W(w) = \int_{-\infty}^{\infty} f_{W,X}(w,x)\,dx$$
$$= \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)\,dx$$

Illustrative Examples (1/4)

- **Example 4.13.** Let X and Y be independent and have PMFs given by

$$p_X(x)=\begin{cases}1/3, & \text{if } x=1,2,3,\\ 0, & \text{otherwise.}\end{cases} \qquad p_Y(y)=\begin{cases}1/2, & \text{if } y=0,\\ 1/3, & \text{if } y=1,\\ 1/6, & \text{if } y=2,\\ 0, & \text{otherwise.}\end{cases}$$

  - Calculate the PMF of $W=X+Y$ by convolution.

We know that the range of possible value of W are integers from the range [1, 5]

$$\begin{aligned}p_W(1)&=\sum_x p_X(x)p_Y(1-x)\\&=p_X(1)p_Y(0)\\&=1/3\cdot 1/2=1/6\end{aligned}$$

$$\begin{aligned}p_W(2)&=\sum_x p_X(x)p_Y(2-x)\\&=p_X(1)p_Y(1)+p_X(2)p_Y(0)\\&=1/3\cdot 1/3+1/3\cdot 1/2\\&=1/9+1/6=5/18\end{aligned}$$

$$\begin{aligned}p_W(3)&=\sum_x p_X(x)p_Y(3-x)\\&=p_X(1)p_Y(2)+p_X(2)p_Y(1)+p_X(3)p_Y(0)\\&=1/3\cdot 1/6+1/3\cdot 1/3+1/3\cdot 1/2\\&=1/18+1/9+1/6=1/3\end{aligned}$$

$$\begin{aligned}p_W(4)&=\sum_x p_X(x)p_Y(4-x)\\&=p_X(2)p_Y(2)+p_X(3)p_Y(1)\\&=1/3\cdot 1/6+1/3\cdot 1/3\\&=1/18+1/9=1/6\end{aligned}$$

$$\begin{aligned}p_W(5)&=\sum_x p_X(x)p_Y(5-x)\\&=p_X(3)p_Y(2)\\&=1/3\cdot 1/6\\&=1/18\end{aligned}$$

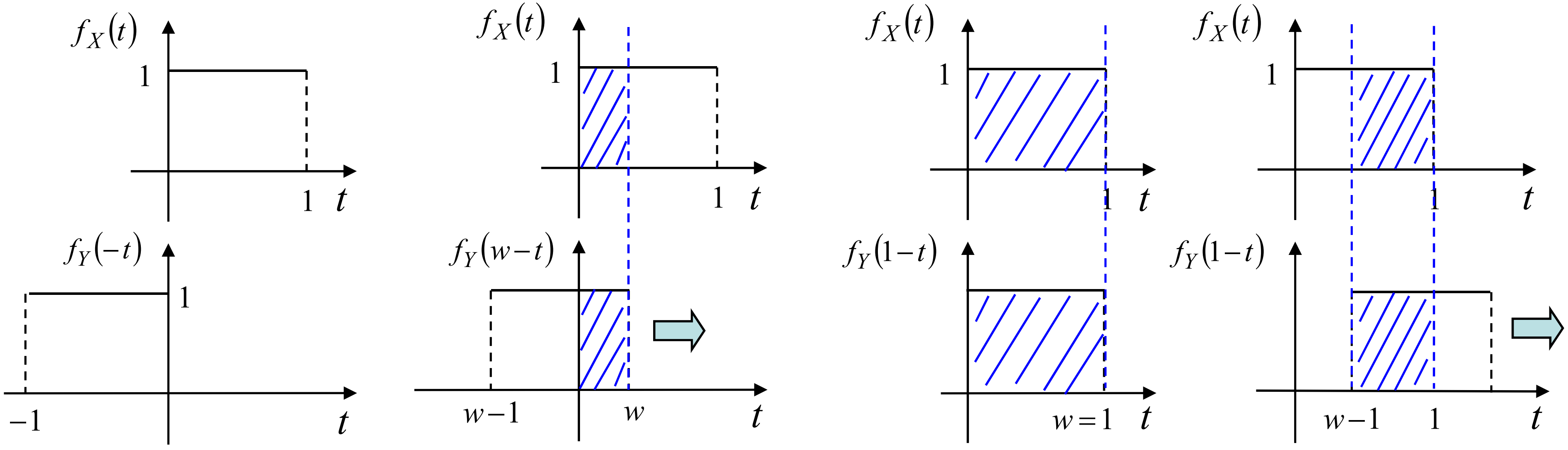
Illustrative Examples (2/4)

- **Example 4.14.** The random variables X and Y are independent and uniformly distributed in the interval [0, 1]. The PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_X(t) f_Y(w-t)\,dt$$

We know that the range of possible value of W are in the range [0, 2]

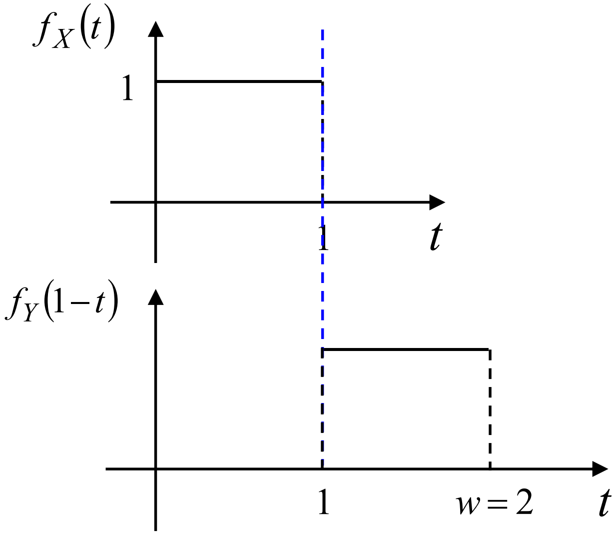
(i) $w = 0$ (ii) $0 < w < 1$ (iii) $w = 1$ (iv) $1 < w < 2$



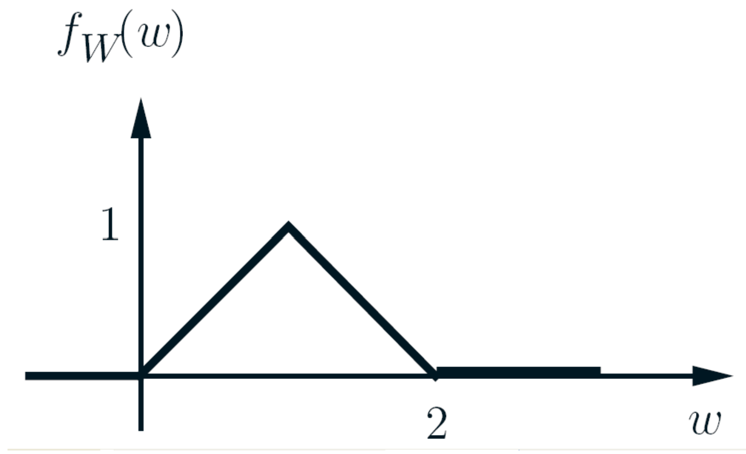
$f_W(0) = \int_0^0 f_X(t) f_Y(-t)\,dt = 0$ $f_W(w) = \int_0^w f_X(t) f_Y(w-t)\,dt = w$ $f_W(1) = \int_0^1 f_X(t) f_Y(1-t)\,dt = 1$ $f_W(w) = \int_{w-1}^1 f_X(t) f_Y(1-t)\,dt = 2 - w$

Illustrative Examples (3/4)

(iv) $w=2$



$$f_W(w)=\int_1^1 f_X(t) f_Y(1-t)\,dt=0$$

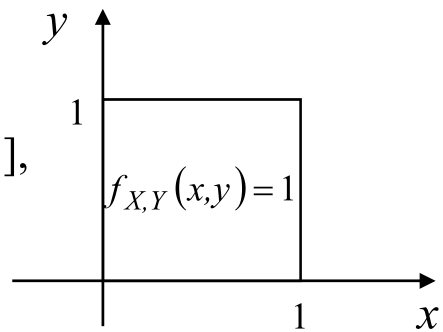


$$\therefore \quad f_W(w)=\begin{cases} w, & \text{if } 0 \le w \le 1 \\ 2-w, & \text{if } 1 \le w \le 2 \\ 0, & \text{otherwise} \end{cases}$$

or as $f_W(w)=\begin{cases} \min\{1, w\} - \max\{0, \mathrm{w}-1\}, & 0 \le w \le 2 \\ 0, & \text{otherwise} \end{cases}$ shown in textbook

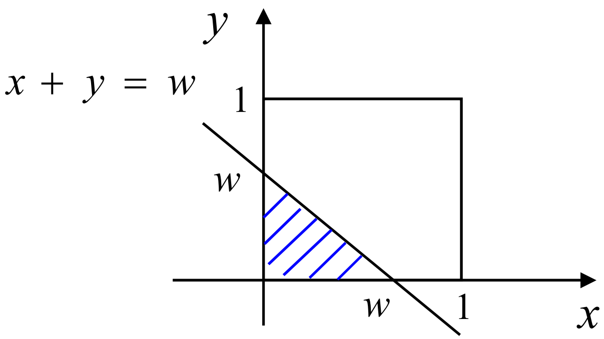
Illustrative Examples (4/4)

- Or, we can use the “**Derived Distribution**” method previously introduced in **Section 3.6**



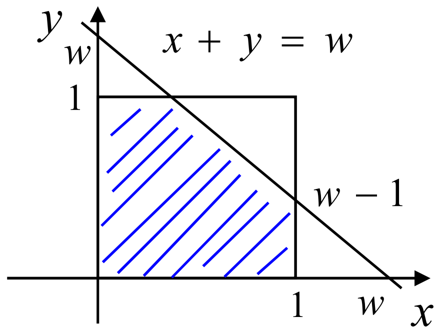
Since X and Y are indepent random varibles uniformly distribute d in $[0,1]$, we have their joint PDF $f_{X,Y}(x,y)=f_X(x)f_Y(y)=1$, for $0 \le x,y \le 1$

(i) $0 \le w \le 1$



$$F_W(w)=\mathbf{P}(W \le w)=\mathbf{P}(X+Y \le w)$$
$$=\frac{1}{2}w^2$$
$$\Rightarrow f_W(w)=w$$

(ii) $1 \le w \le 2$

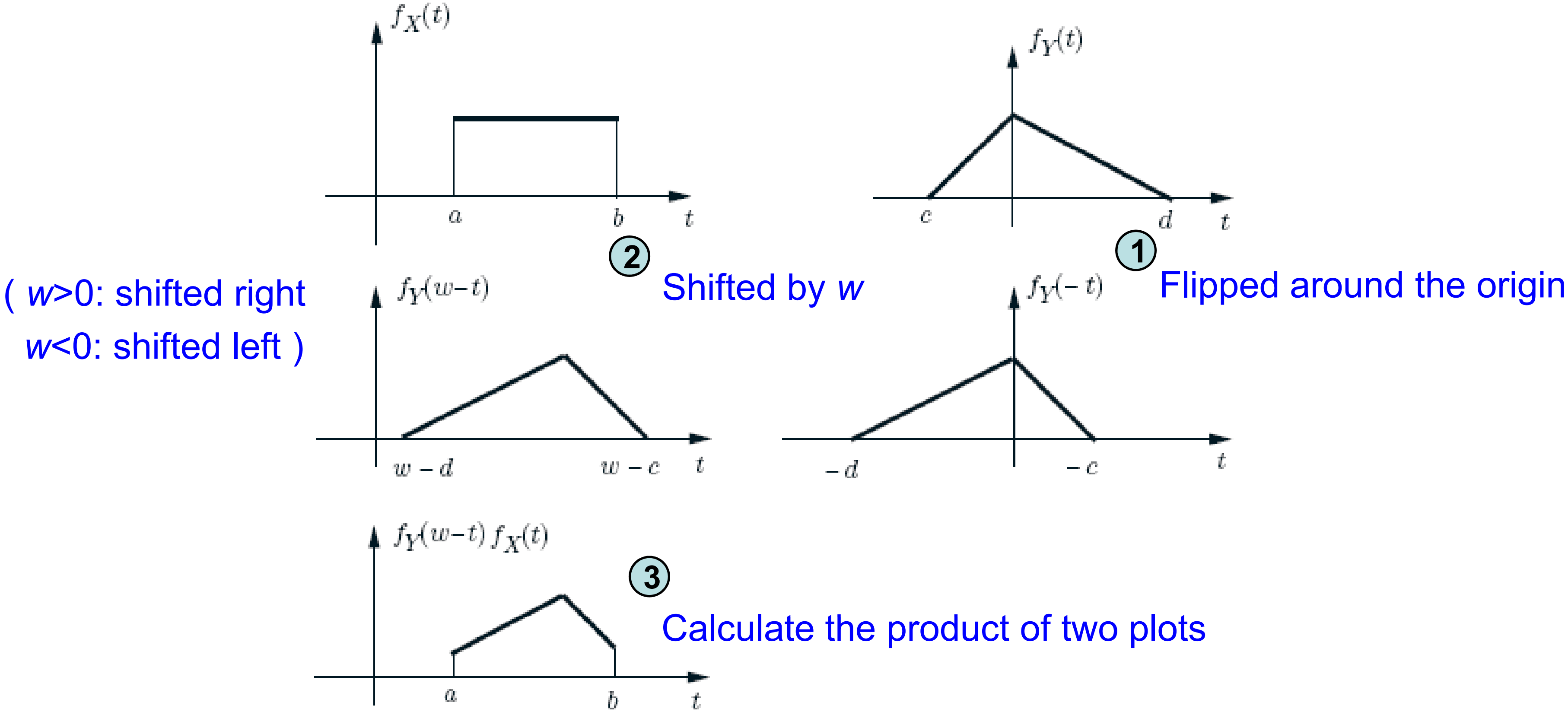


$$F_W(w)=\mathbf{P}(W \le w)=\mathbf{P}(X+Y \le w)$$
$$=1-\frac{1}{2}(2-w)^2$$
$$\Rightarrow f_W(w)=2-w$$

$$\therefore f_W(w)=\begin{cases} w, & \text{if } 0 \le w \le 1 \\ 2-w, & \text{if } 1 \le w \le 2 \\ 0, & \text{otherwise} \end{cases}$$

Graphical Calculation of Convolutions

- **Figure 4.4.** Illustration of the convolution calculation. For the value of W under consideration, $f_W(w)$ is equal to the integral of the function shown in the last plot.



Revisit: Conditional Expectation and Variance

- Goal: To introduce two useful probability laws
  - **Law of Iterated Expectations**

$$\mathbf{E}\big[\mathbf{E}[X|Y]\big] = \mathbf{E}[X]$$

  - **Law of Total Variance**

$$\mathrm{var}(X) = \mathbf{E}\big[\mathrm{var}(X|Y)\big] + \mathrm{var}\big(\mathbf{E}[X|Y]\big)$$

More on Conditional Expectation

- Recall that the conditional expectation $\mathbf{E}\left[X \middle| Y=y\right]$ is defined by

$$\mathbf{E}\left[X \middle| Y=y\right] = \sum_x x \cdot p_{X|Y}\left(x \middle| y\right), \qquad \text{(If } X \text{ is discrete)}$$

and

$$\mathbf{E}\left[X \middle| Y=y\right] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}\left(x \middle| y\right) dx. \quad \text{(If } X \text{ is continuous)}$$

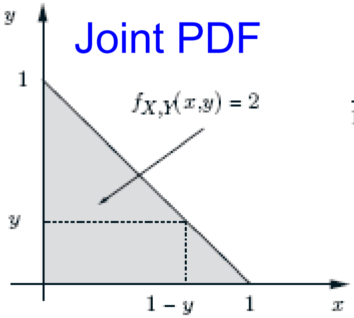
- $\mathbf{E}\left[X \middle| Y=y\right]$ in fact can be viewed as a function of Y, because its value depends on the value y of Y
  - Is $\mathbf{E}\left[X \middle| Y\right]$ a random variable ?
  - What is the expected value of $\mathbf{E}\left[X \middle| Y\right]$?
    - Note also that the expectation of a function $g(Y)$ of Y

$$\mathbf{E}\left[g(Y)\right] = \begin{cases} \sum_y g(y) p_Y(y), & \text{if } Y \text{ is discrete} \\ \int_{-\infty}^{\infty} g(y) f_Y(y) dy, & \text{if } Y \text{ is continuous} \end{cases}$$

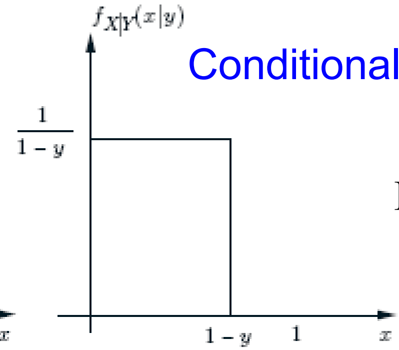
An Illustrative Example (1/2)

- **Example 4.15.** Let the random variables X and Y have a joint PDF which is equal to 2 for (x, y) belonging to the triangle indicated below and zero evervwhere else.

$$X + Y \le 1$$
$$0 \le X \le 1$$
$$0 \le Y \le 1$$



(a)



(b)

$$\mathbf{E}[X|Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)dx.$$

  - What's the value of $\mathbf{E}[X|Y = y]$?

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$$
$$= \int_0^{1-y} f_{X,Y}(x, y)dx \ \ (\because X + Y \le 1)$$
$$= \int_0^{1-y} 2dx = 2(1-y), \qquad 0 \le y \le 1$$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{1-y}$$

for $0 \le x \le 1-y$ where $0 \le y \le 1$

$$\therefore \mathbf{E}[X|Y = y] = \int_0^{1-y} x \cdot \frac{1}{1-y} dx$$
$$= \frac{1}{2(1-y)} \cdot x^2 \Big|_0^{1-y} = \frac{1-y}{2}$$

(a linear function of Y)

An Illustrative Example (2/2)

- We saw that $\mathbf{E}[X|Y=y]=(1-y)/2$. Hence, $\mathbf{E}[X|Y]$ is the random variable $(1-Y)/2$:

$$\mathbf{E}[X|Y]=\frac{(1-Y)}{2}$$

- The expectation of $\mathbf{E}[X|Y]$

$$\mathbf{E}[\mathbf{E}[X|Y]]=\int_{-\infty}^{\infty}\mathbf{E}[X|Y=y]f_Y(y)dy=\mathbf{E}[X]$$

Total Expectation Theorem

For this problem, we thus have

$$\mathbf{E}[X]=\mathbf{E}[\mathbf{E}[X|Y]]=\mathbf{E}[(1-Y)/2]=(1-\mathbf{E}[Y])/2$$

$$\begin{aligned}\mathbf{E}[Y]&=\int_0^1 y\cdot f_Y(y)dy\\&=\int_0^1 y\cdot 2(1-y)dy\\&=y^2-(2/3)y^3\Big|_0^1\\&=1/3\end{aligned}$$

$$\therefore\ \mathbf{E}[X]=(1-\mathbf{E}[Y])/2=1/3$$

Law of Iterated Expectations

$$\mathbf{E}\big[\mathbf{E}[X|Y]\big] = \mathbf{E}[X]$$

$$\mathbf{E}\big[\mathbf{E}[X|Y]\big] = \begin{cases} \sum_y \mathbf{E}[X|Y=y]\,p_Y(y), & \text{(If } Y \text{ is discrete)} \\ \int_{-\infty}^{\infty} \mathbf{E}[X|Y=y]\,f_Y(y)\,dy. & \text{(If } Y \text{ is continuous)} \end{cases}$$

An Illustrative Example (1/2)

- **Example 4.16.** We start with a stick of length l. We break it at a point which is chosen randomly and uniformly over its length, and keep the piece that contains the left end of the stick. We then repeat the same process on the stick that we were left with.
  - What is the expected length of the stick that we are left with, after breaking twice?

Let Y be the length of the stick after we break for the first time.

Let X be the length after the second time.

$$f_Y(y) = \begin{cases} \frac{1}{l}, & \text{for } 0 \le y \le l \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad f_{X|Y}(x|y) = \begin{cases} \frac{1}{y}, & \text{for } 0 \le x \le y \\ 0, & \text{otherwise} \end{cases}$$

uniformly distributed

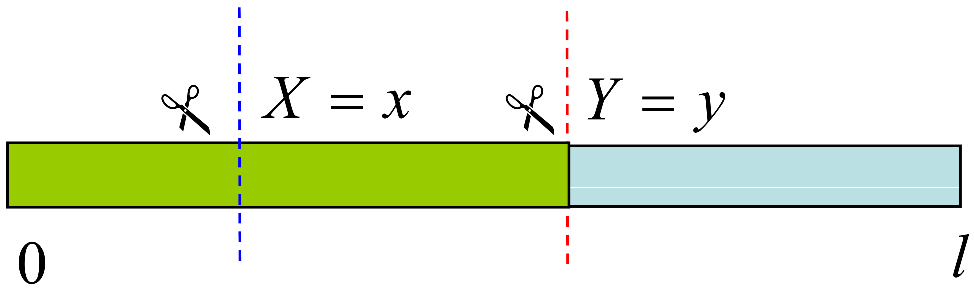
uniformly distributed

An Illustrative Example (2/2)

– By the **Law of Iterated Expectations**, we have

$$
\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}\big[\mathbf{E}[X \mid Y]\big] \\
&= \int_0^l \mathbf{E}[X \mid Y = y] f_y(y)\,dy = \int_0^l \left[\int_0^y x f_{X|Y}(x \mid y)\,dx\right] f_y(y)\,dy \\
&= \int_0^l \left[\int_0^y x \frac{1}{y}\,dx\right] \tfrac{1}{l}\,dy = \int_0^l \left[\frac{1}{y} \cdot \frac{x^2}{2}\Big|_0^y\right] \tfrac{1}{l}\,dy \\
&= \tfrac{1}{l} \cdot \int_0^l \frac{y}{2}\,dy = \tfrac{1}{l} \cdot \frac{y^2}{4}\Big|_0^l \\
&= \frac{l}{4}
\end{aligned}
$$

Note that $\mathbf{E}[X \mid Y = y] = \frac{y}{2}$



Averaging by Section (1/3)

- Averaging by section can be viewed as a special case of the law of iterated expectations

- **Example 4.17. Averaging Quiz Scores by Section.**
  - A class has n students and the quiz score of student i is x_i. The average quiz score is

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i$$

  - If students are divided into k disjoint subsets $A_1, A_2, \ldots, A_k$, the average score in section s is

$$m_s = \frac{1}{n_s}\sum_{x_i \in A_s} x_i$$

Averaging by Section (2/3)

- **Example 4.17. (cont.)**
  - The average score of over the whole class can be computed by taking a weighted average of the average score m_s of each class s, while the weight given to section s is proportional to the number of students in that section

$$\sum_{s=1}^{k} \frac{n_s}{n} m_s = \sum_{s=1}^{k} \frac{n_s}{n} \cdot \frac{1}{n_s} \sum_{x_i \in A_s} x_i$$

$$= \frac{1}{n} \sum_{s=1}^{k} \sum_{x_i \in A_s} x_i$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$= m$$

Averaging by Section (3/3)

- **Example 4.17. (cont.)**
  - **Its relationship with the law of iterated expectations**
    - Two random variable defined
      - X : quiz score of a student (or outcome)
        - » Each student (or outcome) is uniformly distributed
      - Y : section of a student $Y \in \{1, \ldots, k\}$

$$\Rightarrow \mathbf{E}[X] = m \qquad (?)$$

$$\mathbf{E}[X \mid Y = s] = \frac{1}{n_s} \sum_{i \in A_s} x_i = m_s \qquad (?)$$

$$\because P(Y = s) = \frac{n_s}{n} \qquad (?)$$

$$\therefore m = \mathbf{E}[X] = \mathbf{E}\big[\mathbf{E}[X \mid Y]\big] = \sum_{s=1}^{k} \mathbf{E}[X \mid Y = s] P(Y = s)$$

$$= \sum_{s=1}^{k} m_s \cdot \frac{n_s}{n}$$

More on Conditional Variance

- Recall that the conditional variance of X , given $Y=y$, is defined by

$$\mathrm{var}(X|Y=y)=\mathbf{E}\left[\left(X-\mathbf{E}[X|Y=y]\right)^2 \middle| Y=y\right]$$

- $\mathrm{var}(X|Y)$ in fact can be viewed as a function of Y , because its value $\mathrm{var}(X|Y=y)$ depends on the value y of Y
  - Is $\mathrm{var}(X|Y)$ a random variable ?
  - What is the expected value of $\mathrm{var}(X|Y)$?

Note that $\mathbf{E}[\mathrm{var}(X|Y)] \neq \mathrm{var}(X)$

Law of Total Variance

- The expectation of the conditional variance $\mathrm{var}(X|Y)$ is related to the unconditional variance $\mathrm{var}(X)$

$$\mathrm{var}(X) = \mathbf{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbf{E}[X|Y])$$

Law of Iterated Expectations

$$\mathbf{E}[X^2] = \mathbf{E}[\mathbf{E}[X^2|Y]]$$
$$\mathbf{E}[X] = \mathbf{E}[\mathbf{E}[X|Y]]$$

$$\mathrm{var}(Z) = \mathbf{E}[Z^2] - (\mathbf{E}[Z])^2$$
$$\Rightarrow \mathbf{E}[Z^2] = \mathrm{var}(Z) + (\mathbf{E}[Z])^2$$

$$\begin{aligned}
\mathrm{var}(X) &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 \\
&= \mathbf{E}[\mathbf{E}[X^2|Y]] - (\mathbf{E}[\mathbf{E}[X|Y]])^2 \\
&= \mathbf{E}[\mathrm{var}(X|Y) + (\mathbf{E}[X|Y])^2] - (\mathbf{E}[\mathbf{E}[X|Y]])^2 \\
&= \mathbf{E}[\mathrm{var}(X|Y)] + \mathbf{E}[(\mathbf{E}[X|Y])^2] - (\mathbf{E}[\mathbf{E}[X|Y]])^2 \\
&= \mathbf{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbf{E}[X|Y])
\end{aligned}$$

Illustrative Examples (1/4)

- **Example 4.16. (continued)** Consider again the problem where we break twice a stick of length , at randomly chosen points, with Y being the length of the stick after the first break and X being the length after the second break
  - Calculate $\mathrm{var}(X)$ using the law of total variance

$$\mathrm{var}(X) = \mathbf{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbf{E}[X|Y])$$

We know that $f_Y(y) = \begin{cases} \frac{1}{l}, & \text{for } 0 \le y \le l \\ 0, & \text{otherwise} \end{cases}$ and $f_{X|Y}(x|y) = \begin{cases} \frac{1}{y}, & \text{for } 0 \le x \le y \\ 0, & \text{otherwise} \end{cases}$

uniformly distributed ... uniformly distributed

We also know that if a random variable Z is uniformly distributed in $[a, b]$, then its variance is

$$\mathrm{var}(Z) = \frac{(b-a)^2}{12}$$

$$\Rightarrow \mathrm{var}(X|Y=y) = \frac{y^2}{12}$$

$\Rightarrow$ "random variable" $\mathrm{var}(X|Y)$ can be expressed as $\frac{Y^2}{12}$ (That is $\mathrm{var}(X|Y) = \frac{Y^2}{12}$)

(a function of Y)

Illustrative Examples (2/4)

$$\mathbf{E}\left[\operatorname{var}(X|Y)\right] = \int_0^l \operatorname{var}(X|Y=y) f_Y(y)\,dy$$

$$= \int_0^l \frac{y^2}{12}\frac{1}{l}\,dy$$

$$= \frac{y^3}{36\cdot l}\Big|_0^l = \frac{l^2}{36}$$

Note that $\mathbf{E}\left[X|Y=y\right] = \frac{y}{2}$ cf. p.14

$\Rightarrow \mathbf{E}\left[X|Y\right] = \frac{Y}{2}$ (a function of Y)

$\Rightarrow \operatorname{var}\left(\mathbf{E}\left[X|Y\right]\right)$

$$= \operatorname{var}\left(\frac{Y}{2}\right) = \frac{1}{4}\operatorname{var}(Y)$$

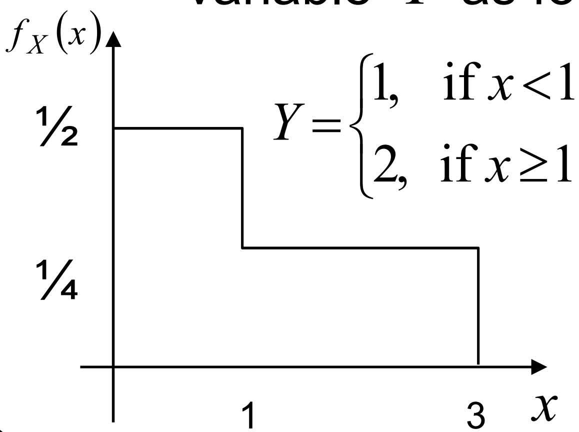
$$= \frac{1}{4}\cdot\frac{l^2}{12} = \frac{l^2}{48}$$ (Y is uniformly distributed)

$$\Rightarrow \quad \therefore \operatorname{var}(X) = \mathbf{E}\left[\operatorname{var}(X|Y)\right] + \operatorname{var}\left(\mathbf{E}\left[X|Y\right]\right)$$

$$= \frac{l^2}{36} + \frac{l^2}{48} = \frac{7\cdot l^2}{144}$$

Illustrative Examples (3/4)

- **Example 4.20.** Computing Variances by Conditioning.
  - Consider a continuous random variable X with the PDF given in the following figure. We define an auxiliary (discrete) random variable Y as follows:



$$Y = \begin{cases} 1, & \text{if } x < 1 \\ 2, & \text{if } x \geq 1 \end{cases}$$

$$\text{var}(X) = \mathbf{E}[\text{var}(X|Y)] + \text{var}(\mathbf{E}[X|Y])$$

We know that if a random variable Z is uniformly distributed in $[a, b]$, then its variance is

$$\text{var}(Z) = \frac{(b-a)^2}{12}$$

1

$$\Rightarrow p_Y(1) = \int_0^1 1/2 dx = 1/2$$

$$p_Y(2) = \int_1^3 1/4 dx = 1/2$$

$$\Rightarrow f_{X|Y}(x|Y=1) = \begin{cases} 1, & \text{for } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{X|Y}(x|Y=2) = \begin{cases} 1/2, & \text{for } 1 \leq x \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

2

$$\text{var}(X|Y=1) = (1-0)^2/12 = 1/12$$

$$\text{var}(X|Y=2) = (3-1)^2/12 = 1/3$$

$$\Rightarrow \mathbf{E}[\text{var}(X|Y)] = \text{var}(X|Y=1)p_Y(1) + \text{var}(X|Y=2)p_Y(2)$$

$$= 1/12 \cdot 1/2 + 1/3 \cdot 1/2 = 5/24$$

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

Illustrative Examples (4/4)

We know that if a random variable Z is uniformly distributed in $[a, b]$, then its mean is

$$E[Z] = \frac{a+b}{2}$$

3

$$\Rightarrow \mathbf{E}[X|Y=1] = (0+1)/2 = 1/2$$

$$\mathbf{E}[X|Y=2] = (1+3)/2 = 2$$

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

$$= \mathbf{E}[X|Y=1]p_Y(1) + \mathbf{E}[X|Y=2]p_Y(2)$$

$$= 1/2 \cdot 1/2 + 2 \cdot 1/2 = 5/4$$

$$\text{var}(\mathbf{E}[X|Y]) =$$

$$(\mathbf{E}[X|Y=1] - \mathbf{E}[\mathbf{E}[X|Y]])^2 p_Y(1)$$

$$+ (\mathbf{E}[X|Y=2] - \mathbf{E}[\mathbf{E}[X|Y]])^2 p_Y(2)$$

$$= (1/2 - 5/4)^2 \cdot 1/2 + (2 - 5/4)^2 \cdot 1/2$$

$$= 9/16$$

Note that for discrete random variable Z

$$\text{var}(Z) = \mathbf{E}\left[(Z - \mathbf{E}(Z))^2\right] = \sum_z (z - \mathbf{E}(Z))^2 p_Z(z)$$

4

$$\therefore \text{var}(X) = \mathbf{E}[\text{var}(X|Y)] + \text{var}(\mathbf{E}[X|Y])$$

$$= 9/16 + 5/24 = 37/48$$

Justification

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx$$

$$= \int_0^1 x \cdot 1/2dx + \int_1^3 x \cdot 1/4dx$$

$$= \frac{1}{4}x^2\Big|_0^1 + \frac{1}{8}x^2\Big|_1^3$$

$$= 5/4$$

$$\text{var}(x) = \int_{-\infty}^{\infty} (x - \mathbf{E}[X])^2 \cdot f_X(x)dx$$

$$= \int_0^1 (x - 5/4)^2 \cdot 1/2dx + \int_1^3 (x - 5/4)^2 \cdot 1/4dx$$

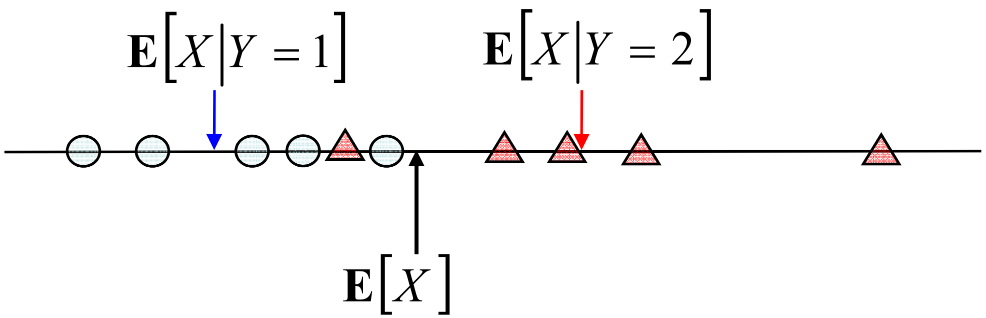
$$= \frac{1}{6}(x - 5/4)^3\Big|_0^1 + \frac{1}{12}(x - 5/4)^3\Big|_1^3$$

$$= \frac{1}{6}\left((-1/4)^3 - (-5/4)^3\right) + \frac{1}{12}\left((7/4)^3 - (-1/4)^3\right)$$

$$= \frac{1}{6} \cdot \frac{124}{64} + \frac{1}{12} \cdot \frac{344}{64} = \frac{37}{48}$$

Averaging by Section

- For a two-section (or two-cluster) problem



○ : $x_i \in$ section 1
△ : $x_i \in$ section 2

$$\mathbf{E}[X] = \mathbf{E}[\mathbf{E}[X|Y]] = \sum_s \mathbf{E}[X|Y=s]P(Y=s)$$

$$\mathrm{var}(X) = \mathbf{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbf{E}[X|Y])$$

These two measures have been widely used for linear discriminant analysis (LDA)

average variability within individual sections

variability of $\mathbf{E}[X|Y]$ (the outcome means of individual sections)

Also called “within cluster” variation

Also called “between cluster” variation

Properties of Conditional Expectation and Variance

- $\mathbf{E}[X \mid Y = y]$ is a number, whose value depends on y.
- $\mathbf{E}[X \mid Y]$ is a function of the random variable Y, hence a random variable. Its experimental value is $\mathbf{E}[X \mid Y = y]$ whenever the experimental value of Y is y.
- $\mathbf{E}\big[\mathbf{E}[X \mid Y]\big] = \mathbf{E}[X]$ (law of iterated expectations).
- $\text{var}(X \mid Y)$ is a random variable whose experimental value is $\text{var}(X \mid Y = y)$, whenever the experimental value of Y is y.
- $\text{var}(X) = \mathbf{E}\big[\text{var}(X \mid Y)\big] + \text{var}\big(\mathbf{E}[X \mid Y]\big)$.

Recitation

- SECTION 4.2 Convolutions
  - Problems 11, 12
- SECTION 4.3 More on Conditional Expectation and Variance
  - Problems 15, 16, 17