2023-2
CB1501012-062

Probability and Statistics

Lecture 11.1: Conditional distribution

Sangryul Jeon
School of Computer Science and Engineering
srjeonn@pusan.ac.kr

Agenda

1. Discrete conditional distribution
2. Continuous conditional distribution
3. Conditional expectation
4. Law of total expectation

Discrete conditional distribution

Recall the definition of the conditional probability of events *E* and *F*

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(EF)}{P(F)}$$

Discrete conditional distribution

Recall the definition of the conditional probability of events *E* and *F*

$$\boldsymbol{P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(EF)}{P(F)}}$$

Joint PDF of X and Y

Conditional PDF of Y

$$p_{Y|X=x}(y) = \frac{p_{XY}(x, y)}{p_X(x)}$$

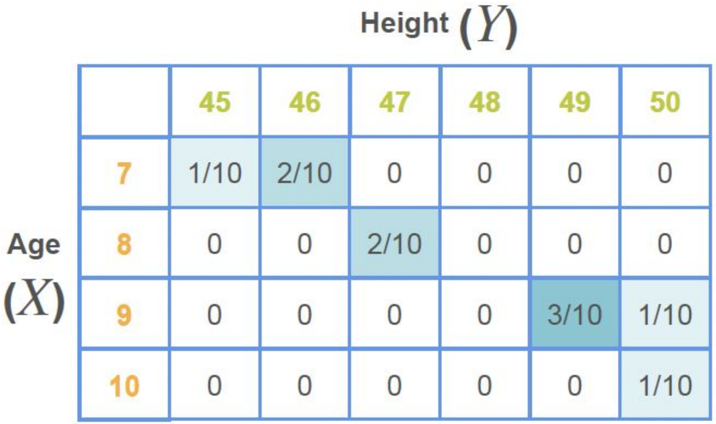
Marginal distribution of X

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50



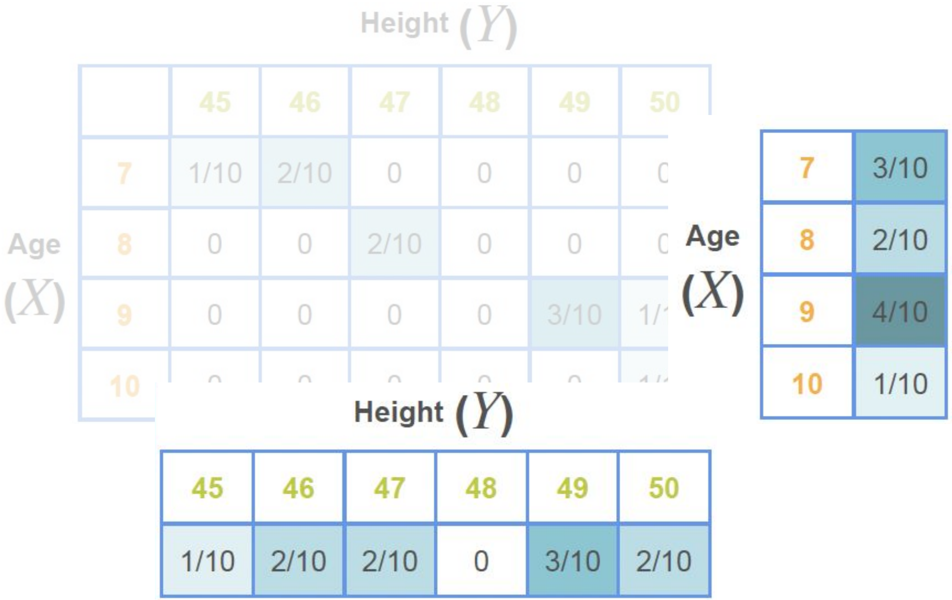
Height (Y)

		45	46	47	48	49	50
Age (X)	7	1/10	2/10	0	0	0	0
	8	0	0	2/10	0	0	0
	9	0	0	0	0	3/10	1/10
	10	0	0	0	0	0	1/10

A quick check-in:
Marginal distributions?

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50



Height (Y)

Age (X)

	45	46	47	48	49	50
7	1/10	2/10	0	0	0	
8	0	0	2/10	0	0	
9	0	0	0	0	3/10	
10						

Age (X)

7	3/10
8	2/10
9	4/10
10	1/10

Height (Y)

45	46	47	48	49	50
1/10	2/10	2/10	0	3/10	2/10

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50

Height (Y)

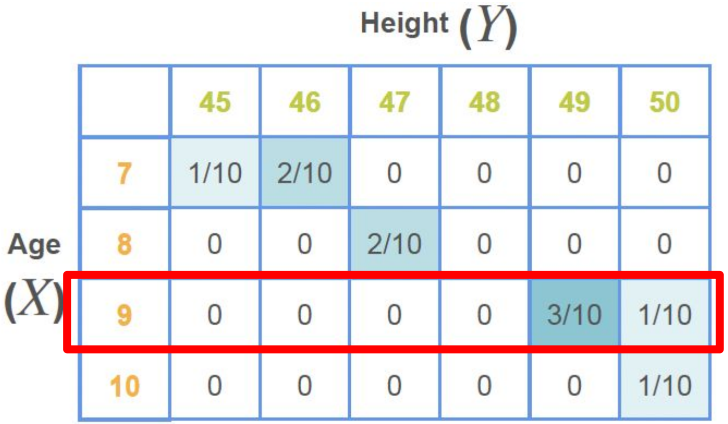
Age (X)

	45	46	47	48	49	50
7	1/10	2/10	0	0	0	0
8	0	0	2/10	0	0	0
9	0	0	0	0	3/10	1/10
10	0	0	0	0	0	1/10

If age=9, what is the distribution across the height variable?

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50



Height (Y)

Age (X)

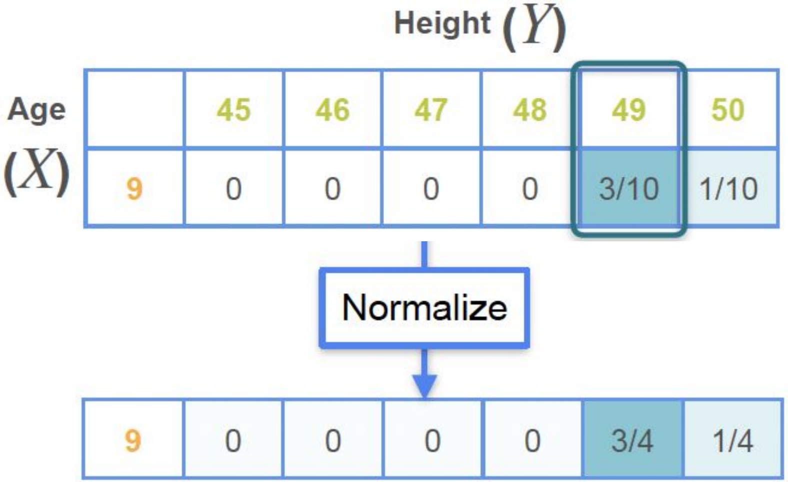
	45	**46**	**47**	**48**	**49**	**50**
7	1/10	2/10	0	0	0	0
8	0	0	2/10	0	0	0
9	0	0	0	0	3/10	1/10
10	0	0	0	0	0	1/10

If age=9, what is the distribution across the height variable?

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y \,|\, X = 9)$$

Discrete conditional distribution

Age (years):	7	7	7	8	8	9	9	9	9	10
Height (in):	45	46	46	47	47	49	49	49	50	50



Height (Y)

Age (X)

	45	46	47	48	49	50
9	0	0	0	0	3/10	1/10

Normalize

9	0	0	0	0	3/4	1/4

$$p_{Y|X=9}(y) = \mathbf{P}(Y = y \mid X = 9)$$

$$P(Y = 49 \mid X = 9) = \frac{\mathbf{P}(X = 9, Y = 49)}{P(X = 9)}$$

Row sum

$$\mathbf{P}(Y = 49 \mid X = 9) = \frac{3/10}{4/10} = \frac{3}{4}$$

Discrete conditional distribution

Die 1:	1/6	1/6	1/6	1/6	1/6	1/6
Die 2:	1/6	1/6	1/6	1/6	1/6	1/6

Discrete conditional distribution

Die 1:	1/6	1/6	1/6	1/6	1/6	1/6
Die 2:	1/6	1/6	1/6	1/6	1/6	1/6

Y

		1	2	3	4	5	6
	1	1/36	1/36	1/36	1/36	1/36	1/36
	2	1/36	1/36	1/36	1/36	1/36	1/36
X	**3**	1/36	1/36	1/36	1/36	1/36	1/36
	4	1/36	1/36	1/36	1/36	1/36	1/36
	5	1/36	1/36	1/36	1/36	1/36	1/36
	6	1/36	1/36	1/36	1/36	1/36	1/36

Discrete conditional distribution

Die 1:	1/6	1/6	1/6	1/6	1/6	1/6
Die 2:	1/6	1/6	1/6	1/6	1/6	1/6

Y

		1	2	3	4	5	6
	1	1/36	1/36	1/36	1/36	1/36	1/36
	2	1/36	1/36	1/36	1/36	1/36	1/36
X	3	1/36	1/36	1/36	1/36	1/36	1/36
	4	1/36	1/36	1/36	1/36	1/36	1/36
	5	1/36	1/36	1/36	1/36	1/36	1/36
	6	1/36	1/36	1/36	1/36	1/36	1/36

$$p_{Y|X=4}(y = 1) = \frac{p_{XY}(x = 4, y = 1)}{p_X(x = 4)}$$

Discrete conditional distribution

Die 1:	1/6	1/6	1/6	1/6	1/6	1/6
Die 2:	1/6	1/6	1/6	1/6	1/6	1/6

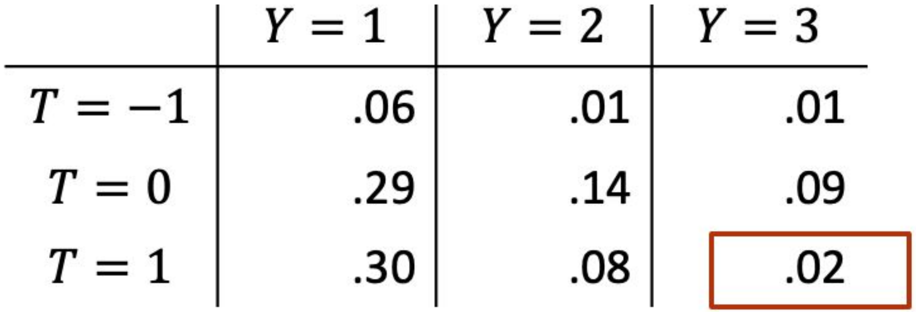
Y

		1	**2**	**3**	**4**	**5**	**6**	**Sum**
	1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
X	**3**	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	6	1/36	1/36	1/36	1/36	1/36	1/36	1/6

$$p_{Y|X=4}(y = 1) = \frac{p_{XY}(x = 4, y = 1)}{p_X(x = 4)} = \frac{1/36}{1/6} = \frac{1}{6}$$

Discrete conditional distribution

Consider we have the following joint PMF, where Y represents the year of the students and T represents the time each student responds



	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$	.06	.01	.01
$T = 0$	.29	.14	.09
$T = 1$	.30	.08	$\boxed{.02}$

$\boldsymbol{P(Y = 3, T = 1)}$

Discrete conditional distribution

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$	.06	.01	.01
$T = 0$	.29	.14	.09
$T = 1$	.30	.08	.02

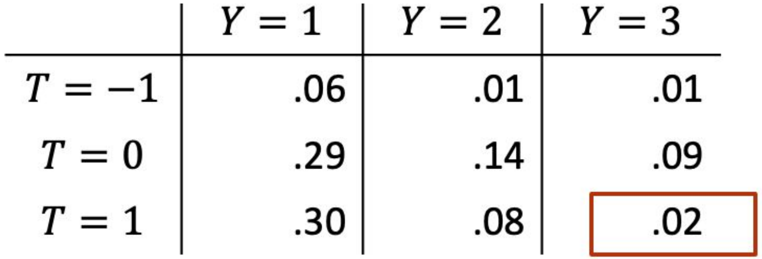
Q1. The below are conditional PMF for (A) $\boldsymbol{P(Y = y|T = t)}$ or (B) $\boldsymbol{P(T = t|Y = y)}$

Which is which?

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$	.09	.04	.08
$T = 0$	.45	.61	.75
$T = 1$	.46	.35	.17

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$	.75	.125	??
$T = 0$	.56	.27	.17
$T = 1$	.75	.2	.05

Discrete conditional distribution

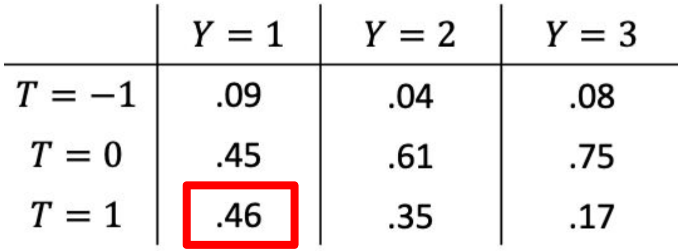


	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.06	.01	.01
$T=0$	.29	.14	.09
$T=1$	.30	.08	.02

Q1. The below are conditional PMF for (A) $\boldsymbol{P(Y=y|T=t)}$ or (B) $\boldsymbol{P(T=t|Y=y)}$

Which is which?

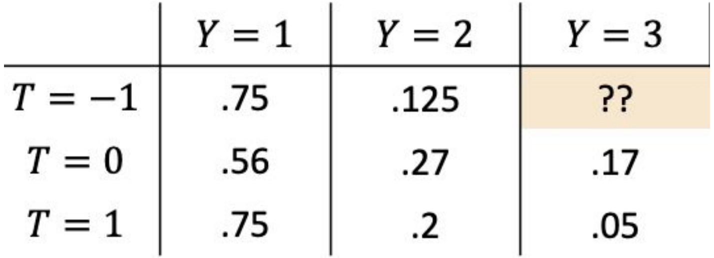
(B) $\boldsymbol{P(T=t|Y=y)}$



	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.09	.04	.08
$T=0$	.45	.61	.75
$T=1$	.46	.35	.17

$\boldsymbol{0.3/(.06+0.29+0.3)}$

(A) $\boldsymbol{P(Y=y|T=t)}$



	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.75	.125	??
$T=0$	.56	.27	.17
$T=1$	.75	.2	.05

Discrete conditional distribution

	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.06	.01	.01
$T=0$	.29	.14	.09
$T=1$	.30	.08	.02

Q2. What is the missing probability?

(B) $P(T=t|Y=y)$

	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.09	.04	.08
$T=0$	.45	.61	.75
$T=1$	.46	.35	.17

(A) $P(Y=y|T=t)$

	$Y=1$	$Y=2$	$Y=3$
$T=-1$	.75	.125	??
$T=0$	.56	.27	.17
$T=1$	.75	.2	.05

Discrete conditional distribution

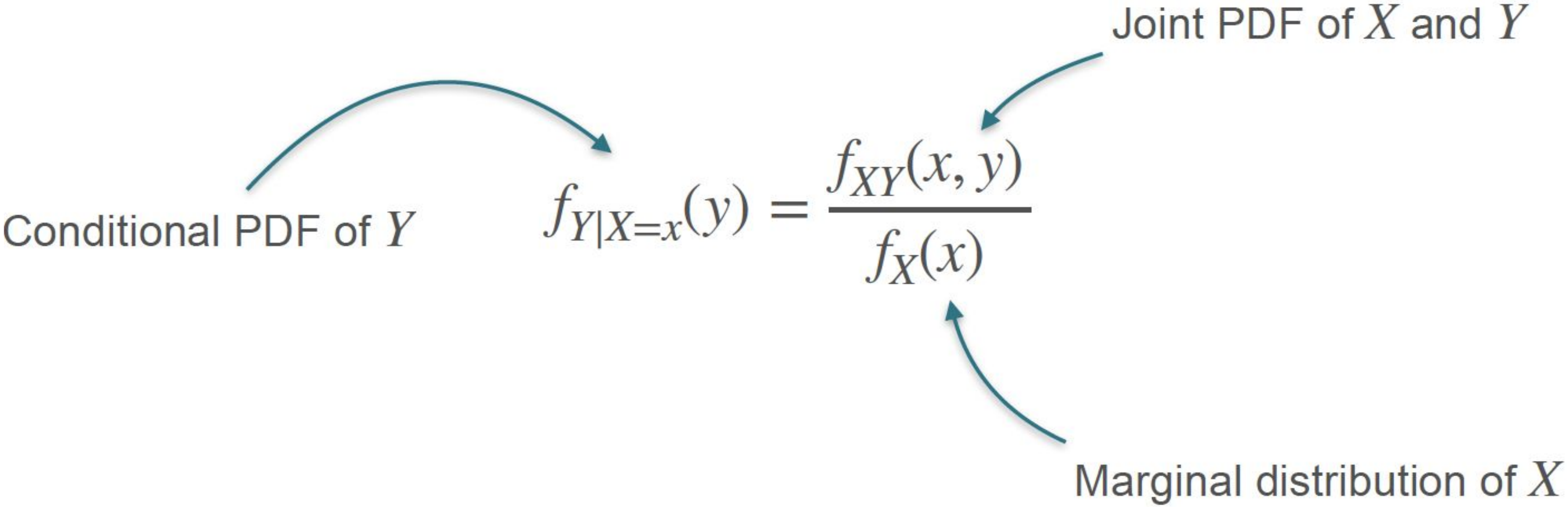
Joint PDF of X and Y

Conditional PDF of Y

$$p_{Y|X=x}(y) = \frac{p_{XY}(x,y)}{p_X(x)}$$

Marginal distribution of X

Continuous conditional distribution



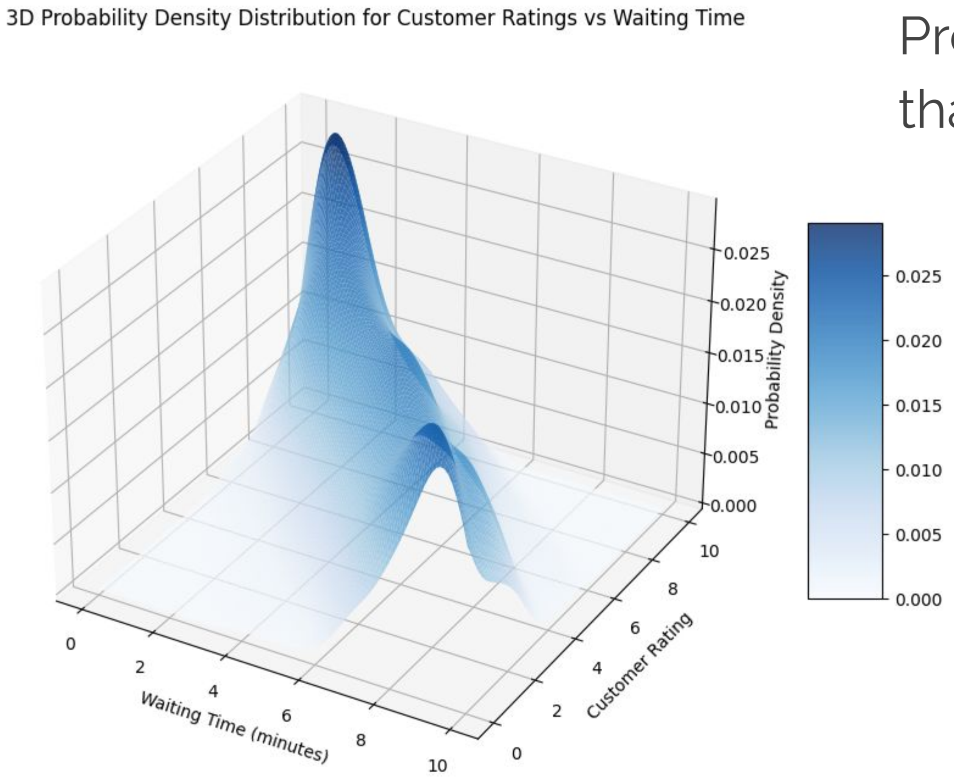
Conditional PDF of Y

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

Joint PDF of X and Y

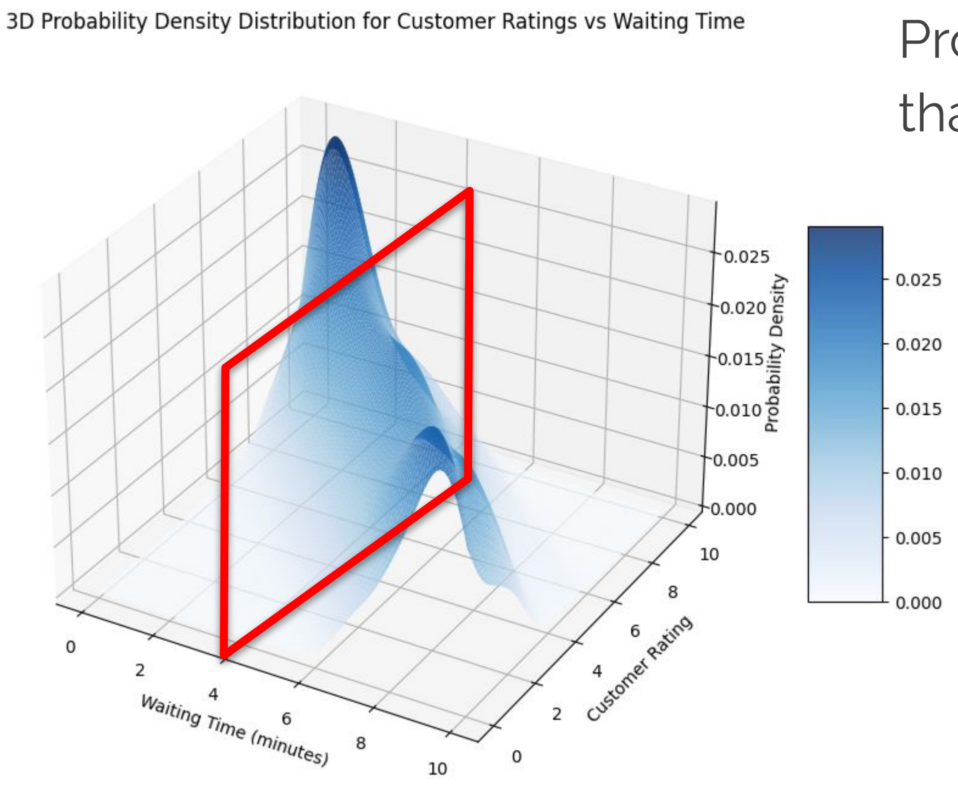
Marginal distribution of X

Continuous conditional distribution

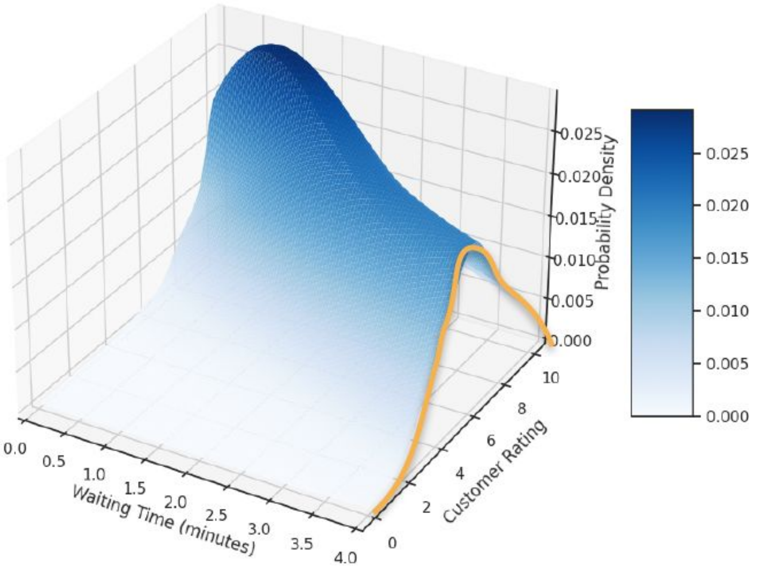


Probability distribution for rating given that waiting time was 4 minutes

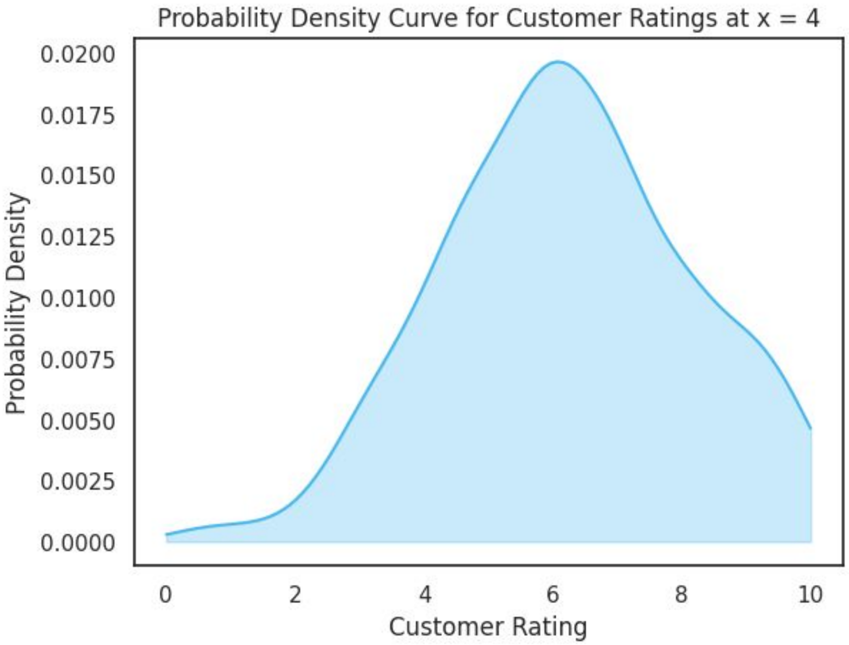
Continuous conditional distribution



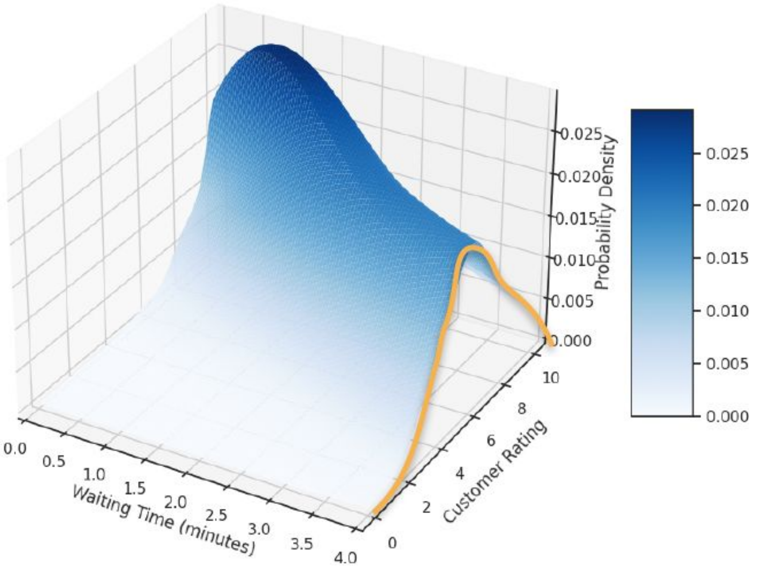
Probability distribution for rating given that waiting time was 4 minutes



Continuous conditional distribution



Probability distribution for rating given that waiting time was 4 minutes



Conditional expectation

Recall the definition of expectation on discrete RV X

$$\mathbb{E}\left[X\right] = \sum_{x \in \Omega_X} x \mathbb{P}\left(X = x\right) = \sum_{x \in \Omega_X} x p_X\left(x\right)$$

Conditional expectation

Recall the definition of expectation on discrete RV X

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x\mathbb{P}(X = x) = \sum_{x \in \Omega_X} x p_X(x)$$

Conditional expectation of X, given knowledge that $Y = y$

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in \Omega_X} x\mathbb{P}(X = x \mid Y = y) = \sum_{x \in \Omega_X} x p_{X,Y}(x \mid y)$$

Conditional expectation

Recall the definition of expectation on discrete RV X

$$\mathbb{E}\left[X\right] = \sum_{x \in \Omega_X} x \mathbb{P}\left(X = x\right) = \sum_{x \in \Omega_X} x p_X\left(x\right)$$

Conditional expectation of X, given knowledge that $Y = y$

$$\mathbb{E}\left[X \mid Y = y\right] = \boxed{\sum_{x \in \Omega_X}} x \mathbb{P}\left(X = x \mid Y = y\right) = \boxed{\sum_{x \in \Omega_X}} x p_{X,Y}\left(x \mid y\right)$$

We are still summing over x and not y

Conditional expectation

Roll two 6-sided dice. Let a random variable S to be a summation of the values of two dices, i.e. $D_1 + D_2$ hat is the expectation of S given the value of second dice is 6?

Conditional expectation

Roll two 6-sided dice. Let a random variable S to be a summation of the values of two dices, i.e. $D_1 + D_2$ hat is the expectation of S given the value of second dice is 6?

$$E[S|D_2 = 6] = \sum_{x=7}^{12} xP(S = x|D_2 = 6)$$

$$= \frac{1}{6}(7 + 8 + 9 + 10 + 11 + 12) \qquad = \frac{57}{6} = 9.5$$

Conditional expectation

More generally,

$$\mathbb{E}\left[X \mid Y = y\right] = \sum_{x \in \Omega_X} x \mathbb{P}\left(X = x \mid Y = y\right) = \sum_{x \in \Omega_X} x p_{X,Y}(x \mid y)$$

$$\mathbb{E}\left[g(X) \mid Y = y\right] = \sum_{x \in \Omega_X} g(x) p_{X|Y}(x \mid y)$$

$$= \int_{-\infty}^{\infty} g(x) f_{X|Y}(x \mid y) dx$$

Conditional expectation

Roll two 6-sided dice. Let a random variable S to be a summation of the values of two dices, i.e. $D_1 + D_2$ hat is the expectation of S given the value of second dice is 6?

$$E[S|D_2 = 6] = E[D_1 + 6|D_2 = 6] \quad = \sum_{d_1} (d_1 + 6) P(D_1 = d_1 | D_2 = 6)$$

$$= \sum_{d_1} d_1 P(D_1 = d_1) + 6 \sum_{d_1} P(D_1 = d_1)$$

$$= E[D_1] + 6 = 3.5 + 6 = 9.5$$

Law of total expectation

Recall the definition of LTE on discrete RV X with independent RVs H and T (head and tail of coin flipping)

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\,\mathbb{P}(H) + \mathbb{E}[X \mid T]\,\mathbb{P}(T)$$

Then, If Y is discrete

$$\mathbb{E}[g(X)] = \sum_{y \in \Omega_Y} \mathbb{E}[g(X) \mid Y = y]\, p_Y(y)$$

If Y is continuous

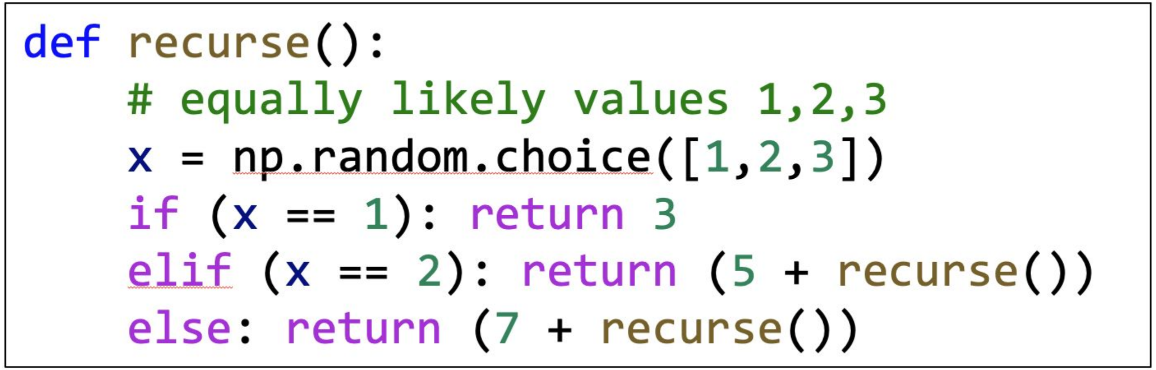
$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} \mathbb{E}[g(X) \mid Y = y]\, f_Y(y) dy$$

Law of total expectation

Proof)

$$\begin{aligned}
\sum_{y\in\Omega_Y} \mathbb{E}\left[g(X) \mid Y = y\right] p_Y(y) &= \sum_{y\in\Omega_Y} \left(\sum_{x\in\Omega_X} g(x) p_{X|Y}(x \mid y) \right) p_Y(y) && \text{def of conditional expectation} \\
&= \sum_{x\in\Omega_X} \sum_{y\in\Omega_Y} g(x) p_{X|Y}(x \mid y) p_Y(y) && \text{swap sums} \\
&= \sum_{x\in\Omega_X} g(x) \sum_{y\in\Omega_Y} p_{X,Y}(x, y) && \text{def of conditional pmf} \\
&= \sum_{x\in\Omega_X} g(x) p_X(x) && \text{def of marginal pmf} \\
&= \mathbb{E}\left[g(X)\right] && \text{def of expectation}
\end{aligned}$$

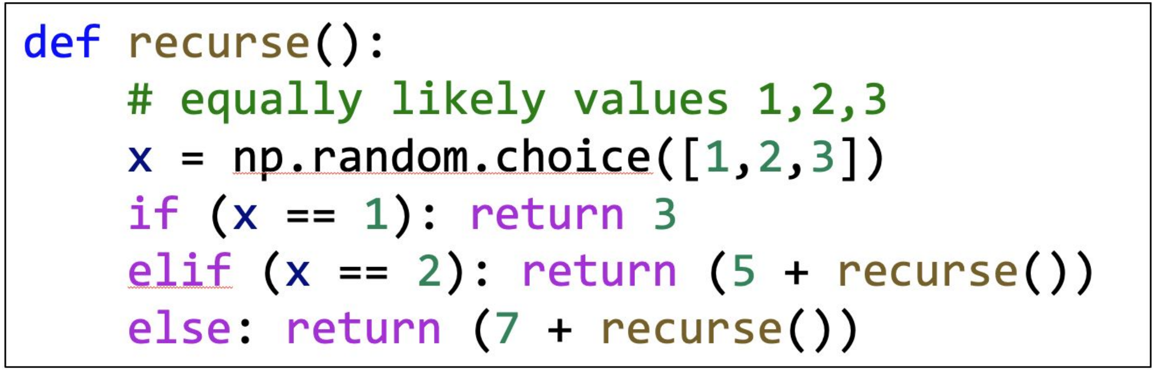
Law of total expectation



```
def recurse():
    # equally likely values 1,2,3
    x = np.random.choice([1,2,3])
    if (x == 1): return 3
    elif (x == 2): return (5 + recurse())
    else: return (7 + recurse())
```

Let Y = return value of recurse(). What is $E[Y]$?

Law of total expectation

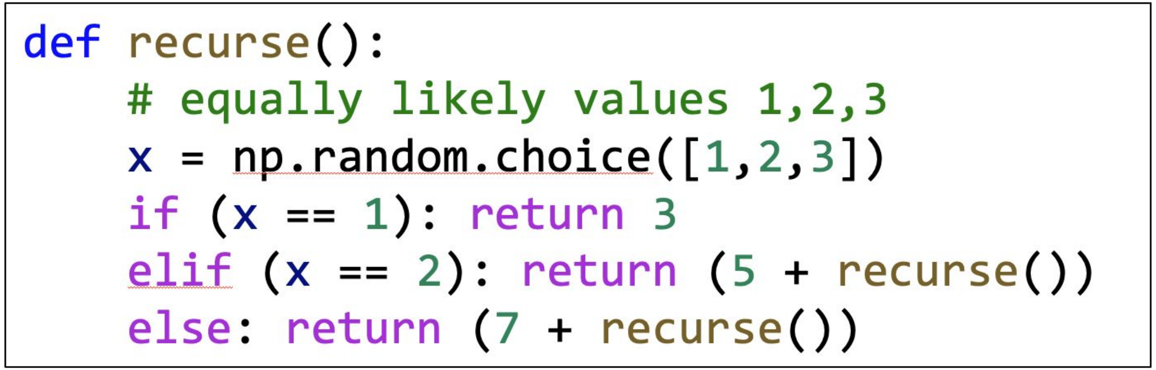


```
def recurse():
    # equally likely values 1,2,3
    x = np.random.choice([1,2,3])
    if (x == 1): return 3
    elif (x == 2): return (5 + recurse())
    else: return (7 + recurse())
```

Let Y = return value of recurse(). What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$

Law of total expectation



```
def recurse():
    # equally likely values 1,2,3
    x = np.random.choice([1,2,3])
    if (x == 1): return 3
    elif (x == 2): return (5 + recurse())
    else: return (7 + recurse())
```

Let Y = return value of recurse(). What is $E[Y]$?

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

$$E[Y|X=1] = 3 \qquad E[Y|X=2] = E[5+Y] \qquad E[Y|X=3] = E[7+Y]$$

$$\mathrm{E[Y]} = \frac{1}{3}\left(15 + 2E[Y]\right) \qquad \mathrm{E[Y]} = 15$$

Law of total expectation

Suppose we get some uniformly random decimal number X from [0, 1]. We keep drawing uniform random numbers until we get a value less than our initial value. What is the expected number of draws until this happens?

Law of total expectation

Suppose we get some uniformly random decimal number X from [0, 1]. We keep drawing uniform random numbers until we get a value less than our initial value. What is the expected number of draws until this happens?

X = uniformly sampled number from [0, 1].

T = the number of draws.

What is $E[T]$?

Law of total expectation

X = uniformly sampled number from [0, 1].

T = the number of draws.

What is $E[T]$?

$$\mathbb{P}(T = t \mid X = x) = (1 - x)^{t-1} x$$

Geometric RV!

$$\begin{aligned}\mathbb{P}(T = t) &= \int_0^1 \mathbb{P}(T = t \mid X = x)\, f_X(x) dx \\ &= \int_0^1 (1 - x)^{t-1} x \cdot 1 dx = \cdots = \frac{1}{t(t+1)}\end{aligned}$$

def of LTE

$$\mathbb{E}[T] = \sum_{t=1}^{\infty} t p_T(t) = \sum_{t=1}^{\infty} t \frac{1}{t(t+1)} = \sum_{t=1}^{\infty} \frac{1}{t+1} = \infty$$

def of expectation