# STAT 24400 Lecture 6
# Section 3.5 Conditional Distributions

Yibi Huang
Department of Statistics
University of Chicago

# Conditional Distributions of Discrete Random Variables

# Example 1 — Gas Station (Revisit)

A gas station has both **self-service** and **full-service** islands, each with a single regular unleaded pump with 2 hoses.

$$X = \text{the \# of hoses in use on the self-service island, and}$$
$$Y = \text{the \# of hoses in use on the full-service island}$$

The joint PMF of $X$ and $Y$:

| | $p(x,y)$ | $Y$ (full-service) 0 | 1 | 2 |
|---|---|---|---|---|
| $X$ | 0 | 0.10 | 0.04 | 0.02 |
| self- | 1 | 0.08 | 0.20 | 0.06 |
| service | 2 | 0.06 | 0.14 | 0.30 |

What is $\mathrm{P}(Y = 1 \mid X = 2)$?

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | Row Sum $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |

| | | $Y$ | | | Row Sum |
|---|---|---|---|---|---|
| | $p(x,y)$ | 0 | 1 | 2 | $p_X(x)$ |
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |

By the definition of conditional probability,

$$\mathrm{P}(Y = 1 \mid X = 2) = \frac{\mathrm{P}(X = 2, Y = 1)}{\mathrm{P}(X = 2)} = \frac{p(2,1)}{p_X(2)} = \frac{0.14}{0.50} = 0.28.$$

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | Row Sum $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |

By the definition of conditional probability,

$$\mathrm{P}(Y = 1 \mid X = 2) = \frac{\mathrm{P}(X = 2, Y = 1)}{\mathrm{P}(X = 2)} = \frac{p(2,1)}{p_X(2)} = \frac{0.14}{0.50} = 0.28.$$

The conditional PMF of $Y$ given $X = 2$ is

$$p_{Y|X}(y \mid x = 2) = \frac{\mathrm{P}(X = 2, Y = y)}{\mathrm{P}(X = 2)} = \frac{p(2,y)}{p_X(2)}$$

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_{Y\|X}(y \mid x = 2)$ | $\frac{0.06}{0.50} = 0.12$ | $\frac{0.14}{0.50} = 0.28$ | $\frac{0.30}{0.50} = 0.60$ |

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |

Similarly, the conditional PMF of $Y$ given $X = 0$ is

$$p_{Y|X}(y \mid x = 0) = \frac{\mathrm{P}(X = 0, Y = y)}{\mathrm{P}(X = 0)} = \frac{p(0, y)}{p_X(0)}$$

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_{Y\|X}(y \mid x = 0)$ | $\frac{0.10}{0.16} = 0.625$ | $\frac{0.04}{0.16} = 0.25$ | $\frac{0.02}{0.16} = 0.125$ |

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |

Similarly, the conditional PMF of $Y$ given $X = 0$ is

$$p_{Y|X}(y \mid x = 0) = \frac{\mathrm{P}(X = 0, Y = y)}{\mathrm{P}(X = 0)} = \frac{p(0, y)}{p_X(0)}$$

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_{Y\|X}(y \mid x = 0)$ | $\frac{0.10}{0.16} = 0.625$ | $\frac{0.04}{0.16} = 0.25$ | $\frac{0.02}{0.16} = 0.125$ |

and the conditional PMF of $Y$ given $X = 1$ is

$$p_{Y|X}(y \mid x = 1) = \frac{\mathrm{P}(X = 1, Y = y)}{\mathrm{P}(X = 1)} = \frac{p(1, y)}{p_X(1)}$$

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_{Y\|X}(y \mid x = 1)$ | $\frac{0.08}{0.34} \approx 0.235$ | $\frac{0.20}{0.34} \approx 0.588$ | $\frac{0.06}{0.34} \approx 0.176$ |

# Conditional Distributions

Suppose $X$ & $Y$ are two discrete r.v.'s with joint PMF $p(x, y)$ and marginal PMF's $p_X(x)$ and $p_Y(y)$ respectively

The **conditional PMF** for $Y$ given $X = x$ is

$$p_{Y|X}(y \mid x) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(X = x)} = \frac{p(x, y)}{p_X(x)},$$

The **conditional PMF** for $X$ given $Y = y$ is

$$p_{X|Y}(x \mid y) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)} = \frac{p(x, y)}{p_Y(y)},$$

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |
| | $p_Y(y)$ | 0.24 | 0.38 | 0.38 | |

| | $p(x,y)$ | $Y$ 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |
| | $p_Y(y)$ | 0.24 | 0.38 | 0.38 | |

| | $p(y \mid x)$ | $Y$ 0 | 1 | 2 | row sum |
|---|---|---|---|---|---|
| | 0 | $\frac{0.10}{0.16} = 0.625$ | $\frac{0.04}{0.16} = 0.25$ | $\frac{0.02}{0.16} = 0.125$ | 1 |
| $X$ | 1 | $\frac{0.08}{0.34} \approx 0.235$ | $\frac{0.20}{0.34} \approx 0.588$ | $\frac{0.06}{0.34} \approx 0.176$ | 1 |
| | 2 | $\frac{0.06}{0.50} = 0.12$ | $\frac{0.14}{0.50} = 0.28$ | $\frac{0.30}{0.50} = 0.6$ | 1 |
| | $p_Y(y)$ | 0.24 | 0.38 | 0.38 | |

- Each row is a PMF for $Y$ given some $x$ value
- Observed the row sums of $p_{Y|X}(y \mid x)$ are all 1

| $p(x,y)$ | | $Y$: 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | 0.10 | 0.04 | 0.02 | 0.16 |
| $X$ | 1 | 0.08 | 0.20 | 0.06 | 0.34 |
| | 2 | 0.06 | 0.14 | 0.30 | 0.50 |
| $p_Y(y)$ | | 0.24 | 0.38 | 0.38 | |

| $p(x \mid y)$ | | $Y$: 0 | 1 | 2 | $p_X(x)$ |
|---|---|---|---|---|---|
| | 0 | $\frac{0.10}{0.24} \approx 0.417$ | $\frac{0.04}{0.38} \approx 0.105$ | $\frac{0.02}{0.38} \approx 0.053$ | 0.16 |
| $X$ | 1 | $\frac{0.08}{0.24} \approx 0.333$ | $\frac{0.20}{0.38} \approx 0.526$ | $\frac{0.06}{0.38} \approx 0.158$ | 0.34 |
| | 2 | $\frac{0.06}{0.24} = 0.25$ | $\frac{0.14}{0.38} \approx 0.368$ | $\frac{0.30}{0.38} \approx 0.790$ | 0.50 |
| column sum | | 1 | 1 | 1 | |

- Each column is a PMF for $X$ given some $y$ value
- Observed the column sums $p_{X|Y}(x \mid y)$ are all 1

In summary,

A conditional PMF of $Y$ given $X = x$ is $p_{Y|X}(y \mid x) = \dfrac{p(x,y)}{p_X(x)}$ which satisfies

$$0 \leq p_{Y|X}(y \mid x) \leq 1 \quad \text{and} \quad \sum_y p_{Y|X}(y \mid x) = 1, \quad \text{for all } x.$$

A conditional PMF of $X$ given $Y = y$ is $p_{X|Y}(x \mid y) = \dfrac{p(x,y)}{p_Y(y)}$ which satisfies

$$0 \leq p_{X|Y}(x \mid y) \leq 1 \quad \text{and} \quad \sum_x p_{X|Y}(x \mid y) = 1, \quad \text{for all } y.$$

# Example 2 — Poisson

For independent r.v.'s $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$, recall in L05, we show that

$$T = X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

**Q**: Given $T = X_1 + X_2 = t$, what's the conditional PMF of $X_1$?

# Example 2 — Poisson

For independent r.v.'s $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$, recall in L05, we show that

$$T = X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

**Q**: Given $T = X_1 + X_2 = t$, what's the conditional PMF of $X_1$?

$$\begin{aligned}
\mathrm{P}(X_1 = x \mid T = t) &= \frac{\mathrm{P}(\{X_1 = x\} \cap \{T = t\})}{\mathrm{P}(T = t)} \\
&= \frac{\mathrm{P}(\{X_1 = x\} \cap \{X_2 = t - x\})}{\mathrm{P}(T = t)} \\
&= \frac{e^{-\lambda_1} \lambda_1^x / x! \cdot e^{-\lambda_2} \lambda_2^{t-x} / (t-x)!}{e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^t / t!} \\
&= \binom{t}{x} \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^x \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{t-x}, \ 0 \le x \le t.
\end{aligned}$$

i.e., given $X_1 + X_2 = t$, $X_1 \sim \text{Bin}(t, \frac{\lambda_1}{\lambda_1+\lambda_2})$.

# Conditional Distributions of Continuous Random Variables

# Conditional Distributions of Continuous Random Variables

Suppose $X$ & $Y$ are two discrete r.v.'s with joint PDF $f(x, y)$ and marginal PMF's $f_X(x)$ and $f_Y(y)$ respectively.

The **conditional PDF** for $X$ given $Y = y$ is

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

The **conditional PDF** for $Y$ given $X = x$ is

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

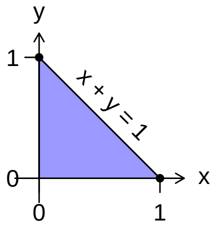
# Example 3 — Deluxe Mixed Nuts

Recall in Lecture 5, the joint PDF for

$$X = \text{ the weight of almonds, and } Y = \text{ the weight of cashews}$$

in a can of mixed nuts is

$$f(x, y) = \begin{cases} 24xy & \text{if } 0 \le x, y \le 1, x + y < 1 \\ 0 & \text{otherwise} \end{cases}$$



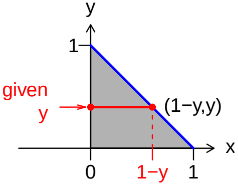

We calculated in L05 the marginal PDF's for $X$ and for $Y$:

$$f_X(x) = 12x(1-x)^2, \quad f_Y(y) = 12y(1-y)^2, \text{ for } 0 \le x, y \le 1.$$

# Conditional Distributions — Mixed Nuts

The conditional PDF $f_{X|Y}(x \mid y)$ of $X$ (almond) given $Y = y$ (cashew) is

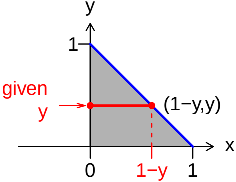
$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$



# Conditional Distributions — Mixed Nuts

The conditional PDF $f_{X|Y}(x \mid y)$ of $X$ (almond) given $Y = y$ (cashew) is

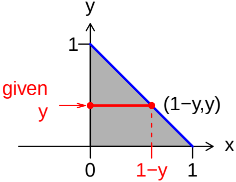
$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{24xy}{12y(1-y)^2}$$



# Conditional Distributions — Mixed Nuts

The conditional PDF $f_{X|Y}(x \mid y)$ of $X$ (almond) given $Y = y$ (cashew) is

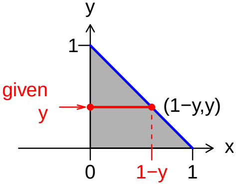
$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{24xy}{12y(1-y)^2}$$
$$= \frac{2x}{(1-y)^2}, \qquad \text{for } 0 \le x \le 1-y\,.$$



# Conditional Distributions — Mixed Nuts

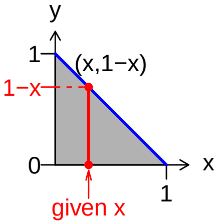
The conditional PDF $f_{X|Y}(x \mid y)$ of $X$ (almond) given $Y = y$ (cashew) is

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{24xy}{12y(1-y)^2}$$
$$= \frac{2x}{(1-y)^2}, \qquad \text{for } 0 \le x \le 1-y\,.$$




Similarly, the conditional PDF $f_{Y|X}(y \mid x)$ of $Y$ (cashew) given $X = x$ (almond) is

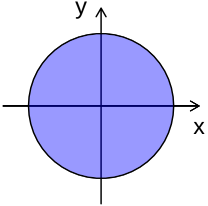
$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{2y}{(1-x)^2}, \quad \text{for } 0 \le y \le 1-x\,.$$



# Example 4: Uniform Disk

$(X, Y)$ is chosen uniformly at random from the unit disk, $\{x^2 + y^2 \le 1\}$. The joint PDF is

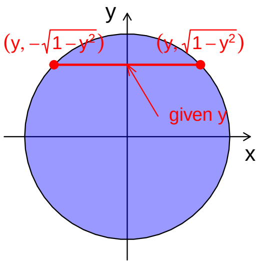
$$f(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \le 1, \\ 0, & \text{otherwise} \end{cases}$$




Marginal PMF of $Y$

$$f_Y(y) = \int_{x=-\infty}^{\infty} f(x, y)\mathrm{d}x = \int_{x=-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}\mathrm{d}x$$

$$= \frac{2}{\pi}\sqrt{1 - y^2} \quad \text{for } -1 \le y \le 1$$




PMF for $X$ is identical $f_X(x) = \dfrac{2}{\pi}\sqrt{1 - x^2}$ for $-1 \le x \le 1$.

## Example 4: Uniform Disk (2)

The Conditional PDF for $X$ given $Y = y$ is

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{1}{\pi}}{\frac{2\sqrt{1-y^2}}{\pi}} = \frac{1}{2\sqrt{1-y^2}}$$

for $-\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}$, which is constant in $x$.

In other words, given $Y = y$,

$$X \sim \text{Uniform}\left(-\sqrt{1-y^2},\ \sqrt{1-y^2}\right).$$

Likewise, given $X = x$, $Y$ is Uniform on $\left(-\sqrt{1-x^2}, \sqrt{1-x^2}\right)$.

Note that

- the *marginal* PDF of $Y$ (or of $X$) is not uniform, but
- the *conditional* PDF of $X \mid Y$ (or of $Y \mid X$) is uniform.

# Example 5

Recall on p.40 of L05 slides, for $X$ and $Y$ w/ the joint PDF

$$f(x, y) = 6xy^2, \quad \text{for } 0 \le x, y \le 1,$$

We found the marginal PDF's of $X$ and of $Y$ to be

$$f_X(x) = 2x,\ 0 < x < 1, \quad \text{and} \quad f_Y(y) = 3y^2,\ 0 < y < 1.$$

The conditional PDF of $y$ given $X = x$ is

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{6xy^2}{2x} = 3y^2, \quad 0 < y < 1.$$

which is exactly the marginal PDF of $Y$.

Recall in L05, we said $X$ and $Y$ are **independent** since
$f(x, y) = 6xy^2 = (2x)(3y^2) = f_X(x)f_Y(y)$ for all $0 \le x, y \le 1$ and
$f(x, y) = 0 = f_X(x)f_Y(y)$ elsewhere.

# Conditional = Marginal, when Independent

What is the conditional distribution of $Y$ given $X = x$ if $X$ and $Y$ are independent?

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

i.e., conditional PDF $Y|X$ is the marginal PDF of $Y$.

# Conditional = Marginal, when Independent

What is the conditional distribution of $Y$ given $X = x$ if $X$ and $Y$ are independent?

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

i.e., conditional PDF $Y|X$ is the marginal PDF of $Y$.

In fact, the following three are equivalent definitions of the independence of $X$ and $Y$

- $f(x, y) = f_X(x) f_Y(y)$ ............................ (joint = product of marginal)
- $f_{Y|X}(y \mid X = x) = f_Y(y)$ .................. (conditional $Y \mid X$ = marginal of $Y$)
- $f_{X|Y}(x \mid Y = y) = f_X(x)$ ................. (conditional $X \mid Y$ = marginal of $X$)

# Conditional = Marginal, when Independent

What is the conditional distribution of $Y$ given $X = x$ if $X$ and $Y$ are independent?

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

i.e., conditional PDF $Y|X$ is the marginal PDF of $Y$.

In fact, the following three are equivalent definitions of the independence of $X$ and $Y$

- $f(x, y) = f_X(x) f_Y(y)$ ............................ (joint = product of marginal)
- $f_{Y|X}(y \mid X = x) = f_Y(y)$ .................. (conditional $Y \mid X$ = marginal of $Y$)
- $f_{X|Y}(x \mid Y = y) = f_X(x)$ ................. (conditional $X \mid Y$ = marginal of $X$)

All the things above apply to joint/conditional/marginal **PMF** for discrete $X, Y$., too.