MATH 321: Mathematical Statistics

Lecture 15: Conditional Distributions

Chapter 4: Bivariate Distributions (4.3)

Introduction

- Oftentimes, two random variables (X, Y) are related. Knowing about the value of X gives us some information about the value of Y, even if it doesn't tell us the value Y exactly (can find $E(Y \mid X = x)$, but not the exact value of $Y \mid X = x$).
- Example: Study hours X and Test grade Y.

  $P(Y > 90 \mid X = 1 \text{ hrs}) < P(Y > 90 \mid X = 5 \text{ hrs})$
- Sometimes, knowledge about X gives us no information about Y. $\Rightarrow$ X & Y are independent

Discrete conditional distributions

Conditional pmf

- Recall the conditional probability of events: $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(X=x, Y=y)}{P(X=x)}$
- Events in a conditional distribution.
  - Suppose that X and Y are discrete random variables. The conditional event of $Y = y$ given $X = x$ is

    $\{\overset{B}{Y = y} \mid \overset{A}{X = x}\}$ events

    where $X = x$ is the conditioning event (i.e. the given event),

    and $Y = y$ is the event of interest (i.e. the event whose probability we want to know).

- Definition: Let (X, Y) be a discrete bivariate random vector with joint pmf $f(x, y)$ and marginal pmfs $f_X(x)$ and $f_Y(y)$.

  (a) For any x such that $P(X = x) = f_X(x) > 0$ $(x \in \mathcal{X})$, the **conditional pmf of Y given that $X = x$** is the function of y denoted by $f(y \mid x)$ and defined by

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

  (b) For any y such that $P(Y = y) = f_Y(y) > 0$ $(y \in \mathcal{Y})$, the **conditional pmf of X given that $Y = y$** is the function of x denoted by $f(x \mid y)$ and defined by

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

Probabilities

- Once we have the conditional pmf, we can find probabilities as expected.

  For $A \subset \mathbb{R}^2$,

  $$P(X \in A \mid Y = y) = \sum_{x \in A} P(X = x \mid Y = y) = \sum_{x \in A} f(x \mid y)$$

  (just flip for $y \mid x$)

- We can also show that the conditional pmf is indeed a valid pmf.

  Proof, need to show:

  1. $f(x \mid y) \geq 0$ for all x.
  2. $\sum_x f(x \mid y) = 1$.

① $f(x,y) \geq 0$ for all y

By definition $f(x|y) = \dfrac{P(X=x, Y=y)}{P(Y=y)} = \dfrac{\geq 0}{> 0} \Rightarrow \overset{\checkmark}{\geq 0}$ overall

② $\sum_x f(x|y) = \sum_x \dfrac{f(x,y)}{f_Y(y)}$ → constant because y is fixed

$= \dfrac{1}{f_Y(y)} \sum_x f(x,y)$

$= \dfrac{1}{\cancel{f_Y(y)}} \cancel{f_Y(y)}$

$\overset{\checkmark}{=} 1$

Examples

1. Interpreting distributions:
   - Let X = GPA and Y = study hours per day.

     If we are given the joint pmf $f(x, y) = P(X = x, Y = y) \longrightarrow$ Probability a student has GPA = x and study hours = y

     then we can find the following:

     i) $f_X(x) = \sum_y f(x,y)$ $\longrightarrow$ Probability student has GPA = x

     ii) $f_Y(y) = \sum_x f(x,y)$ $\longrightarrow$ Probability student has study hours = y

     iii) $f(x \mid y) = f(x,y)/f_Y(y)$ $\longrightarrow$ Probability student has GPA = x given study hours = y

     iv) $f(y \mid x) = f(x,y)/f_X(x)$ $\longrightarrow$ Probability student has study hours = y given GPA = x

2. Define the joint pmf of (X, Y) by:

   $f(0, 10) = f(0, 20) = 2/18, \quad f(1, 10) = f(1, 30) = 3/18,$

   $f(1, 20) = 4/18, \quad \text{and} \quad f(2, 30) = 4/18.$

   (a) Compute the conditional pmf of Y given X for each of the possible values of X.

↳ Joint Pmf table and marginal of X

$y \backslash x$	0	1	2
10	2/18	4/18	0
20	2/18	4/18	0
30	0	3/18	4/18
$f_X(x)$	4/18	10/18	4/18

$$\rightarrow f(y|x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

$$\rightarrow f(y \mid x=0) = \frac{P(X=0, Y=y)}{P(X=0)} = \begin{cases} \frac{2/18}{4/18} = 1/2 & Y=10 \\ \frac{2/18}{4/18} = 1/2 & Y=20 \\ 0 & Y=30 \end{cases}$$

$$\rightarrow f(y \mid x=1) = \frac{P(X=1, Y=y)}{P(X=1)} = \begin{cases} \frac{3/18}{10/18} = 3/10 & Y=10 \\ \frac{4/18}{10/18} = 4/10 & Y=20 \\ \frac{3/18}{10/18} = 3/10 & Y=30 \end{cases}$$

$$\rightarrow f(y \mid x=2) = \frac{P(X=2, Y=y)}{P(X=2)} = \begin{cases} 0 & Y=10 \\ 0 & Y=20 \\ \frac{4/18}{4/18} = 1 & Y=30 \end{cases}$$

   (b) Find $(X = 2, Y > 20)$. $= f(2, 30) = 4/18$

   Joint event

   (c) Find $P(X < 1)$. $= P(X=0) = 4/18$

   marginal event

   (d) Find $P(Y > 10 \mid X = 0)$. $= P(Y = \{20, 30\} \mid X=0) = 1/2$

   conditional event

3. In a previous example, we had the joint pmf

$$f(x, y) = \frac{x+y}{21} \qquad \text{for } x = 1, 2, 3 \text{ and } y = 1, 2.$$

And we found the marginal distributions:

$$f_X(x) = \frac{2x+3}{21} \qquad \text{for } x = 1, 2, 3$$

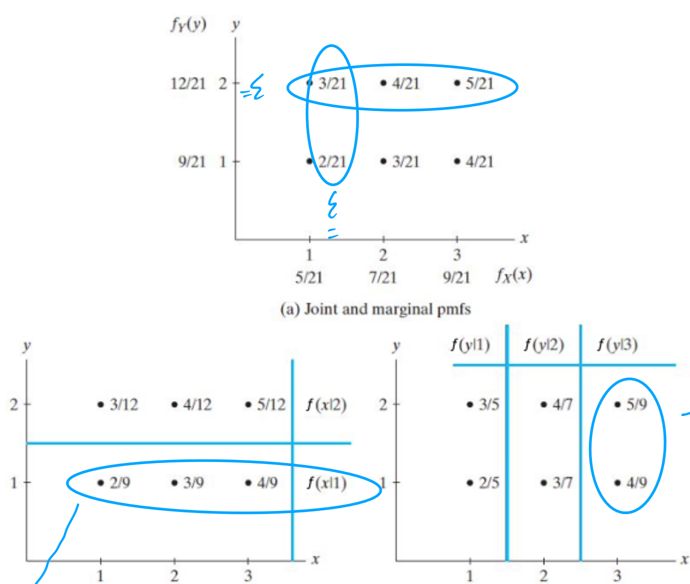
$$f_Y(y) = \frac{3y+6}{21} = \frac{y+2}{7} \qquad \text{for } y = 1, 2$$

Find $f(x \mid y)$ and $f(y \mid x)$.

$\rightarrow f(x|y) = \dfrac{f(x,y)}{f_Y(y)} = \dfrac{(x+y)/21}{(3y+6)/21} = \dfrac{x+y}{3y+6}$ for $x = 1, 2, 3$ when $y = 1, 2$

$\rightarrow f(y|x) = \dfrac{f(x,y)}{f_X(x)} = \dfrac{(x+y)/21}{(2x+3)/21} = \dfrac{x+y}{2x+3}$ for $y = 1, 2$ when $x = 1, 2, 3$

Plots of ranges with corresponding probabilities for all distributions:



(a) Joint and marginal pmfs

(b) Conditional pmfs of X, given y

(c) Conditional pmfs of Y, given x

$f(y|x=3) = \dfrac{3+y}{2(3)+3} = \dfrac{3+y}{9}$, $y = 1, 2$

$f(x|y=1) = \dfrac{x+1}{3(1)+6} = \dfrac{x+1}{9}$, $x = 1, 2, 3$

Conditional random variable

Understanding conditional random variables

- $Y \mid X = x$ is a random variable about Y having the conditional pmf of $f(y \mid x)$.

  The conditional random variables $Y \mid X = 0$ and $Y \mid X = 1$ have different pmfs.
- **The conditional pmf $f(y \mid x)$ is determined by** x **and thus** x **behaves like a parameter** (e.g. Geometric(p)),

Relationship between joint pmf and conditional pmfs

- The following theorem contains the relationship between the joint pmf of X and Y and the two conditional pmfs $f(y \mid x)$ and $f(x \mid y)$.
- Theorem: For bivariate random vector (X, Y) with joint pmf $f(x, y)$ and x and y such that $f_X(x) > 0$ and $f_Y(y) > 0$,

Both ways

Same as general multiplication rule

$$f(x, y) = f_Y(y) \cdot f(x \mid y) = f_X(x) \cdot f(y \mid x)$$

Joint = marginal * conditional

$\to P(B) \cdot P(A \mid B) = \frac{P(A \cap B)}{P(B)} \cdot P(B) \Longrightarrow P(A \cap B) = P(B) \cdot P(A \mid B)$

$\to f_Y(y) \cdot f(x \mid y) = \frac{f(x,y)}{f_Y(y)} \cdot f_Y(y)$

Continuous conditional distributions

Conditional pdf

- If X and Y are continuous random variables, then $P(X = x) = 0$, for every value of x.

  $\Longrightarrow$ Can't use $\frac{f(x,y)}{P(X=x)}$ because it is undefined.

  To define a conditional probability distribution for Y given $X = x$ when X and Y are both continuous is analogous to the discrete case with pdfs replacing pmfs.

- Definition: Let (X, Y) be a continuous bivariate random vector with joint pdf $f(x, y)$ and marginal pmfs $f_X(x)$ and $f_Y(y)$.

  (a) Given x such that $f_X(x) > 0$, $\quad f(y \mid x) = \dfrac{f(x, y)}{f_X(x)}$

  (b) Given y such that $f_Y(y) > 0$, $\quad f(x \mid y) = \dfrac{f(x, y)}{f_Y(y)}$

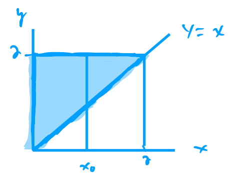
Example

- In a previous example, we had the joint pdf

  $f(x, y) = 1/2 \quad$ for $0 \le x \le y \le 2$.

  And we found the marginal distributions:

  $f_X(x) = (2-x)/2 \quad$ for $0 \le x \le 2 \qquad$ and $\qquad f_Y(y) = y/2 \quad$ for $0 \le y \le 2$



  (a) For $0 \le x < 2$, find the conditional pdf $f(y \mid x)$.

  – NOTE: The range of $Y \mid X = x$ often depends on x. To help, you should draw the range of X and Y just like when finding joint probabilities.

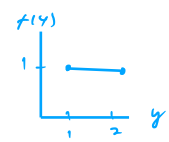
  – For $0 \le x < 2$,

  Need this because have to look at range where $f_X(x) > 0$

  $$f(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{1/2}{(2-x)/2} = \frac{1}{2-x}, \quad x \le y \le 2$$

  Conditioned on $X = x$, we see that $Y \mid X = x \sim$ Uniform$(x, 2)$

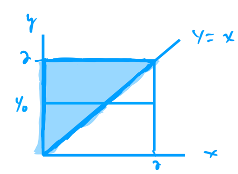
  $\Longrightarrow$ Constant with respect to y



  (b) Find the distribution of $Y \mid X = 1$ $\sim$ Uniform$(1, 2)$

  (we have a specific "parameter" value now).

  (c) For $0 < y \le 2$, find the conditional pdf $f(x \mid y)$.



  – For $0 < y \le 2$,

  $$f(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{1/2}{y/2} = 1/y, \quad 0 \le x \le y$$

  – Given $Y = y$, we see that $X \mid Y = y \sim$ Uniform$(0, y)$

  (d) Find the distribution of $X \mid Y = 1.5$ $\sim$ Uniform$(0, 1.5)$

  [redacted]

  $$f(x \mid y{=}1.5) = \frac{f(x, 1.5)}{f_Y(1.5)} = \frac{1/2}{1.5/2} = \frac{1}{1.5}, \quad 0 \le x \le 1.5$$

  (e) Find the conditional probability that $X \le 1/2 \mid Y = 1.5$.

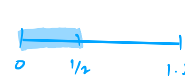
  ↳ Need random variable $X \mid Y = 1.5$

  $$P(X \le 0.5 \mid Y = 1.5) = \int_{-\infty}^{1/2} f(x \mid y{=}1.5)\,dx$$

  $$= \int_0^{1/2} 1/1.5 \; dx$$

  $$= 1/3$$

  OR uniform



  $\frac{1/2}{3/2} = 1/3$

Expected value of a conditional random variable

Conditional expectations and when to use which density

- In addition to their usefulness for calculating probabilities, the conditional pmfs and pdfs can also be used to calculate expected values.

  Just remember that $f(y \mid x)$ as a function of y is a pmf or pdf; so use it in the same way that we have previously used unconditional pmfs or pdfs.

- Suppose, we have $f(x, y)$, $f_X(x)$, $f_Y(y)$, $f(y \mid x)$ and $f(x \mid y)$. What density function should we use to compute the following?

  1. $E(X) =$ $\int x f(x)\, dx$
  2. $E(Y^2) =$ $\int y^2 f(y)\, dy$
  3. $E(Y - Y) =$ $E(0) = 0$
  4. $E(X^2 Y) =$ $\iint x^2 y\, f(x,y)\, dx\, dy$
  5. $E(Y \mid X = 2) =$ $\int y\, f(y \mid x=2)\, dy$
  6. $E(Y^2 \mid X = 3) =$ $\int y^2 f(y \mid x=3)\, dy$
  7. $E(X \mid X = 3) =$ $E(3) = 3$
  8. $E(X + Y^2 \mid Y = 3) =$ $\int (x + 3^2)\, f(x \mid y=3)\, dx$

     OR $= E(x \mid y=3) + E(y^2 \mid y=3) = E(x \mid y=3) + 9$
  9. $E(XY \mid X = 3) =$ $E(3y \mid x=3) = 3E(y \mid x=3) = 3\int y\, f(y \mid x=3)\, dy$

Conditional expected values

- Definition: Let $g(Y)$ be a function of Y, then the **conditional expected value of $g(Y)$ given that $X = x$** is denoted by $E[g(Y) \mid X = x]$ and is given by

$$E[g(Y) \mid x] = \sum g(y) f(y \mid x) \qquad \text{and} \qquad E[g(Y) \mid x] = \int_{-\infty}^{\infty} g(y) f(y \mid x)\, \mathrm{d}y$$

  in the discrete and continuous cases, respectively.

- Conditional mean and variance definitions (assuming X and Y are discrete):
  i) If $g(Y) = Y$, then the **conditional mean of Y given $X = x$** is

  $$E(Y|X=x) = \sum_y y f(y|x) = \mu_{Y|X=x} = \mu_{Y|X}$$

  ii) If $g(Y) = (Y - \mu_{Y|X})^2$, then the **conditional variance of Y given $X = x$** is

  $$E[(Y-\mu_{Y|X})^2|X=x] = \sum_y (y-\mu_{Y|X})^2 f(y|x)$$
  $$= E(Y^2|X=x) - (\underbrace{E(Y|X=x)}_{\mu_{Y|X}})^2 = \sigma^2_{Y|X}$$

Examples

1. In a previous example, we had the joint pmf

   $f(x, y) = \dfrac{x + y}{21}$ for $x = 1, 2, 3$ and $y = 1, 2$.

   $f(x|y=1) = \frac{x+1}{3(1)+6} = \frac{x+1}{9}, \ x=1,2,3$

   And we found the conditional distribution:

   $f(x \mid y) = \dfrac{x + y}{3y + 6}$ for $x = 1, 2, 3$ when $y = 1, 2$.

   (a) Find $\mu_{X|1}$.

   $$\mu_{X|1} = \sum_x x f(x|y=1) = \sum_{x=1}^{3} x\left(\frac{x+1}{9}\right) = \frac{1}{9}[1(2) + 2(3) + 3(4)] = \frac{20}{9}$$

   (b) Find $\sigma^2_{X|1}$.

   $$\rightarrow E(X^2|Y=1) = \sum_x x^2 f(x|y=1) = \sum_{x=1}^{3} x^2\left(\frac{x+1}{9}\right) = \frac{1}{9}[1^2(2) + 2^2(3) + 3^2(4)] = \frac{50}{9}$$

   $$\rightarrow \sigma^2_{X|1} = E(X^2|Y=1) - (\mu_{X|1})^2 = \frac{50}{9} - \left(\frac{20}{9}\right)^2 = \frac{50}{81}$$

2. For $0 < x \leq 1$, the conditional pdf of $Y \mid X = x$ is $f(y \mid x) = \dfrac{2y}{x^2}$ $\quad 0 \leq y \leq x$.

   Note: For this example, the range of __Y__ depends on __x__. So the density as well as the range change when __$X=x$__ is given.

   (a) Find $E(Y \mid X = x)$.

   $$E(Y|X=x) = \int_0^x y f(y|x)\,dy$$
   $$= \int_0^x y\left(\frac{2y}{x^2}\right)dy$$
   $$= \frac{2}{3x^2} y^3 \Big|_0^x$$
   $$= \frac{2}{3}x, \quad 0 < x \leq 1$$

(b) Find the conditional variance $V(Y \mid X = 0.5)$. .

$$V(Y|X=0.5) = E(Y^2|X=0.5) - (E(Y|X=0.5))^2$$
$$= \int_0^{0.5} y^2 \frac{2y}{0.5^2} dy - (\frac{2}{3}x)^2 \quad x \to 0.5$$
$$= 1/8 - (1/3)^2$$
$$= 1/72$$

Understanding conditional expectation

- $E(X)$, $E(Y)$, $E(XY)$ are fixed constants → Center is Not a random variable.
- How about $E(Y \mid X = x)$?

  Let's compare the following two conditional expectations.

$$E(Y \mid X = 1/2) = \int_{-\infty}^{\infty} y\, f(y|1/2)\, dy$$

$$E(Y \mid X = 1) = \int_{-\infty}^{\infty} y\, f(y|1)\, dy$$

- The conditional expectation depends only on the Conditional pdf of $y|x=x$ which is determined by the value of x. Consequently, the conditional expected value, $E(Y \mid X = x)$, is determined by the value of x.

  In other words, as x changes, $E(Y \mid X = x)$ changes. Thus, $E(Y \mid X = x)$ is a function of x.

- What if x is not specified like $E(Y \mid X)$? Then $E(Y \mid X)$ is a function of random variable of X, and thus it is a random variable.

  When x is not specified, replace x by X. Then $E(Y \mid X) =$ $h(X)$

- **Why is conditional expectation important?**
  - **Regression Analysis**. The main purpose of regression analysis is to identify $E(Y|X)$, which explains the mean behavior of Y given X.
  - In regression analysis, we usually assume that Y and X have a **linear relationship**, that is $E(Y \mid X) = \beta_0 + \beta_1 X$.
  - We will study this more later.

Intercept + slope $\cdot X$

$E(Y|X) = \beta_0 + \beta_1 X$