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.

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

Discrete conditional distributions

Conditional pmf

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

    where ____________ is the conditioning event (i.e. the given event),

    and ____________ 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 $\boldsymbol{Y}$ given that $\boldsymbol{X = x}$** is the function of y denoted by $f(y \mid x)$ and defined by

  $$f(y \mid x) = P(Y = y \mid X = x) =$$

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

  $$f(x \mid y) = P(X = x \mid Y = y) =$$

Probabilities

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

  For $A \subset \mathbb{R}^2$,
  $$P(X \in A \mid Y = y) = \sum_{x \in A} P(X = x \mid Y = y) =$$

  (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. $\sum_{x} f(x \mid y) = 1$.

Examples

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

     If we are given the joint pmf $f(x, y) = P(X = x, Y = y) \longrightarrow$

     then we can find the following:

     i) $f_X(x) =$ $\longrightarrow$ Probability student has

     ii) $f_Y(y) =$ $\longrightarrow$ Probability student has

     iii) $f(x \mid y) =$ $\longrightarrow$ Probability student has

     iv) $f(y \mid x) =$ $\longrightarrow$ Probability student has

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.

   (b) Find $(X = 2, Y > 20)$

   ____________ event

   (c) Find $P(X < 1)$

   ____________ event

   (d) Find $P(Y > 10 \mid X = 0)$

   ____________ event

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

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

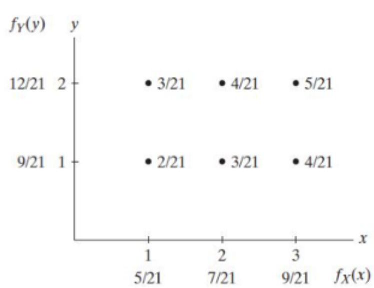
   And we found the marginal distributions:

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

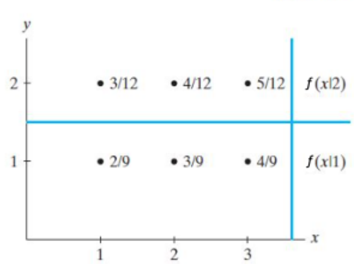
   $$f_Y(y) = \frac{3y+6}{21} = \frac{y+2}{7} \qquad \text{for } y = 1, 2$$

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

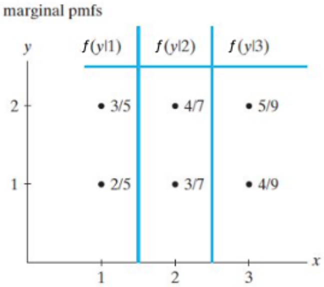
   Plots of ranges with corresponding probabilities for all distributions:



(a) Joint and marginal pmfs



(b) Conditional pmfs of X, given y



(c) Conditional pmfs of Y, given x

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 _____ and thus _____ 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 $f_Y(y) > 0$,

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

Continuous conditional distributions

Conditional pdf

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

  $\Longrightarrow$ Can't use $\dfrac{f(x, y)}{P(X = x)}$ 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 $f_X(x) > 0$, $\qquad f(y \mid x) = \dfrac{f(x, y)}{f_X(x)}$

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

Example

- In a previous example, we had the joint pdf

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

  And we found the marginal distributions:

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

  (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 $0 \le x < 2$,

      $$f(y \mid x) =$$

      Conditioned on $X = x$, we see that $Y \mid X = x \sim$

  (b) Find the distribution of $Y \mid X = 1$

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

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

    - For $0 < y \le 2$,

      $$f(x \mid y) =$$

    - Given $Y = y$, we see that $X \mid Y = y \sim$

  (d) Find the distribution of $X \mid Y = 1.5$

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

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. $E(X) =$
  2. $E(Y^2) =$
  3. $E(Y - Y) =$
  4. $E(X^2 Y) =$
  5. $E(Y \mid X = 2) =$
  6. $E(Y^2 \mid X = 3) =$
  7. $E(X \mid X = 3) =$
  8. $E(X + Y^2 \mid Y = 3) =$
  9. $E(XY \mid X = 3) =$

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 $\boldsymbol{Y}$ given $\boldsymbol{X = x}$** is

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

Examples

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

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

   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 $\mu_{X|1}$.

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

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

   Note: For this example, the range of ______ depends on ______. So the density as well as the range change when ____________ is given.

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

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

Understanding conditional expectation

- $E(X)$, $E(Y)$, $E(XY)$ are ________________ $\rightarrow$ Center is ______________________.
- How about $E(Y \mid X = x)$?

  Let's compare the following two conditional expectations.

$$E(Y \mid X = 1/2) = \int_{-\infty}^{\infty} y \qquad\qquad \mathrm{d}y$$
$$E(Y \mid X = 1) = \int_{-\infty}^{\infty} y \qquad\qquad \mathrm{d}y$$

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

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

- 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 ______________________.

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

- **Why is conditional expectation important?**
  - **Regression Analysis**. The main purpose of regression analysis is to identify ___________, 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.