MATH 321: Mathematical Statistics

Lecture 14: Bivariate Distributions

Chapter 4: Bivariate Distributions (4.1 and 4.4)

Introduction

- In previous chapters, we have discussed probability models and computation of probability for events involving only one random variable. These are called **univariate models**.
- In this chapter, we will discuss models that involve more than one random variable called **multivariate models**.
- Examples:
  - Univariate: The body weights of several people in the population is measured.
  - Multivariate: Temperature, height, and blood pressure, in addition to weight, are measured. These observations on different characteristics could also be modeled as observations on different random variables.
- We need to know how to describe and use probability models that deal with more than one random variable at a time.

  The basic setting for multivariate random variables is the same as those for univariate random variables.
- Definition of a random variable:
  - A random variable is a function from the sample space S to $\mathbb{R}$.

    Univariate:

    Multivariate:
  - We could extend this for an n-dimensional random vector:

– We will mainly discuss **bivariate models**, involving two random variables. With each point in a sample space, we associate an ordered pair of of numbers $(x, y) \in \mathbb{R}^2$, where $\mathbb{R}^2$ denotes the plane $\Longrightarrow$ Possible range is the (x, y) coordinate plane.

Joint probability function for discrete random variables

Example

- An investor owns two assets. They are interested in the value of each of them during one year. It is not enough to know the separate probability distributions, they want to know how the two assets behave together.
- This requires a joint probability distribution for X and Y, which can be specified in a table:

y \ x	90	100	110
0	.05	.27	.18
10	.15	.33	.02

Definition

- The random vector (X, Y) is called a **discrete random vector** if it has only a ________________ number of possible values (i.e. ___________ range).

- Definition: Let (X, Y) be a discrete bivariate random vector. Then the **joint probability mass function (joint pmf)** is defined as $f(x, y) = P(X = x, Y = y)$ for all $(x, y) \in \mathbb{R}^2$ and has properties
  1. $f_{X,Y}(x, y)$ for all x, y.
  2. $\sum \sum f_{X,Y}(x, y) =$

     where the sum is over all value (x, y) that are assigned nonzero probabilities.
  3. Let A be any subset of $\mathbb{R}^2$, then

     $P((X, Y) \in A) = \sum \sum f(x, y)$.

- The joint pmf of (X, Y) completely defines the probability distribution of the random vector (X, Y), just as the pmf of a discrete univariate random variable does.

Examples

1. Consider the experiment of tossing two fair 3-sided die. Let X = sum of two dice and $Y = |$ difference of the two dice $|$.

   (a) Find the sample space and the range of (X, Y).

   (b) Find the joint pmf of (X, Y).

   Recall the **sample point method** for probabilities: $f(x, y) = \frac{\#\text{ outcomes in }(x,y)}{\text{total }\#\text{ outcomes}}$

   (c) Find $P(3Y \geq X)$.

2. Joint probability functions for discrete random variables are often given in tables, but they may also be given in formulas.

   An analyst is studying traffic accidents in two adjacent towns. The random variables X and represents the number of accidents in a day in towns X and Y, respectively. The joint probability function for X and Y is given by:

   $f(x, y) = \frac{e^{-2}}{x!y!}$ for $x = 0, 1, 2, \ldots$ and $y = 0, 1, 2, \ldots$.

   Find $P(X = 1, Y < 2)$.

Marginal distributions for discrete random variables

Motivation and example

- Even if we are considering a probability model for a random vector (X, Y), there may be probabilities of interest that involve only one of the random variables in the vector. For example, $P(X = 2)$.
- $\{X = x\}$ suggests that Y ______________________________ as long as the condition on X is met. This corresponds to the joint event:

  $\{X = x\} =$
- Once we know the joint distribution, it is really easy to find the probabilities for individual values of X and Y.

  Joint pmf table: Marginals =
- Back to the investor example, if they want to know how each individual asset is behaving:

y \ x	90	100	110
0	.05	.27	.18
10	.15	.33	.02

Definition

- Let (X, Y) be a discrete bivariate random vector with joint pmf $f_{X,Y}(x, y)$. Then, the **marginal pmfs** of X and Y, $f_X(x) = P(X = x)$ and $f_Y(y) = P(Y = y)$ are given by

$$f_X(x) = \sum_y f_{X,Y}(x, y) \qquad \text{and} \qquad f_Y(y) = \sum_x f_{X,Y}(x, y)$$

Example

- Given the joint random vector (X, Y), let

$$f(x, y) = \begin{cases} c(x + y) & \text{for } x = 1, 2, 3 \text{ and } y = 1, 2 \\ 0 & \text{elsewhere} \end{cases}$$

  (a) Find c so that $f(x, y)$ is a valid pmf.

(b) Find $f_X(x)$ and $f_Y(y)$ (NOTE: should be functions of ONLY x and ONLY y, respectively).

(c) Find $P(X \leq 2)$.

Probabilities for only one random variable $\Longrightarrow$ Find marginal first.

Summary

- The joint pmf of (X, Y) has more information about the distribution of (X, Y) than the marginal pmfs of X and Y alone.

  This is because it contains information about the relationship between X and Y. And can easily find the marginal distributions from the joint distribution.

Joint and marginal distributions for continuous random variables

Joint distribution definition

- We can also consider random vectors whose components are continuous random variables. The probability distribution of a continuous random vector is usually described using a pdf, as in the univariate case.
- The definitions are really the same except that integrals replace summations.
- Definition: The **joint probability density function (joint pdf)** is a function $f(x, y)$ from $\mathbb{R}^2$ into $\mathbb{R}$ such that
  1. $f_{X,Y}(x, y)$ for all x, y.
  2. $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y\,\mathrm{d}x = 1$ (switch order of integration)
  3. For $A \subset \mathbb{R}^2$

     $$P((X, Y) \in A) = \int\int_A f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int\int_A f(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

Probabilities

- For a continuous random variable X with pdf $f_X(x)$ and $A = [a, b]$:

- Finding $P((X, Y) \in A)$ can be a bit more complicated because it involves double integration.

  For a bivariate continuous random vector (X, Y) with joint pdf $f(x, y)$ and $A = [a, b] \times [c, d]$:

- Examples:
  1. Suppose $f(x, y) = 2 - 1.2x - 0.8y \quad$ for $0 \le x \le 1, 0 \le y \le 1$.
     (a) Verify this is a valid pdf.

     (b) Rectangular region: Find $P(0.50 \le X \le 1, 0.50 \le Y \le 1)$.

(c) More general region:
Find $P(X+Y>1)$, $\qquad f(x,y)=2-1.2x-0.8y \quad$ for $0\le x\le 1, 0\le y\le 1$

(1) Draw range $(\mathcal{X},\mathcal{Y})$ $\qquad$ (2) Draw and shade conditions on (x,y)

(3) Set bounds $\qquad$ (3) Set bounds (again)

Steps

1) Draw the region where the density function is positive (the range of X and Y).
2) Shade the region of the probability we want (make $Y<f(X)$ or $Y>f(X)$).
3) Choose the order of integration and set the bounds of the integrals. Start with the outside integral, then the inside integral.

   Suppose we choose x as the outside integral. Find the interval of x in the shaded region. Then find the interval of y in the shaded region as a function of x

(“moving x”).

2. Suppose $f(x, y) = 3x \quad$ for $0 \leq y \leq x \leq 1$.

   Find $P(0 \leq X \leq 0.5, Y > 0.25)$.

   (1) Draw range $(\mathcal{X}, \mathcal{Y})$

   (2) Draw and shade conditions on (x, y)

   (3) Set bounds

   (3) Set bounds (again)

3. Suppose $f(x, y) = 1 \quad$ for $0 \leq x \leq 1, 0 \leq y \leq 1$.

   Find $P(XY < 0.5)$.

   – NOTE: Constant density $\Longrightarrow$ Uniform distribution (i.e. flat surface)

     Probability of a uniform distribution is just a ________________

     $\mathbb{R}^1$: Prob $= \frac{\text{length of interval of interest}}{\text{length of entire interval}}$

     $\mathbb{R}^2$: Prob $= \frac{\text{Area of interest}}{\text{Total area}}$

Marginal distributions

- Definition: Let $f(x, y)$ be the joint pdf for the bivariate continuous random vector (X, Y). Then the **marginal pdfs** are defined by:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y \qquad \text{and} \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x$$

- Examples:
  1. Continuing previous example 1: $f(x, y) = 2 - 1.2x - 0.8y$ for $0 \le x \le 1, 0 \le y \le 1$.
     (a) Find $f_X(x)$ Should be a function of ONLY x.

     (b) Find $f_Y(y)$. Should be a function of ONLY y.

  2. Suppose $f(x, y) = 1/2$ for $0 \le x \le y \le 2$.
     (a) Find $f_X(x)$.

     (b) Find $f_Y(y)$.

Expected values of functions of random variables

Introduction

- Many practical applications require the study of a function of two or more random variables. For example, for the investor that owns two assets, they may wish to find the distribution of $g(X, Y) = X + Y$, which gives the total value of the two assets.
- Other common functions include (where we also need fancier techniques to find the distribution functions):

  $XY \rightarrow$ Requires a bivariate transformation in the continuous case.

  $\min(X, Y)$ and $\max(X, Y) \rightarrow$ Order statistics.
- For now, we are not going to find distributions of functions of random variable. Rather, we are going focus on expectations of functions of random vectors $g(X, Y)$, which can be computed from the original joint distribution.

  These can be computed just as with univariate random variables.

Notation

- Sometimes it is convenient to replace the symbols X and Y representing random variables by X_1 and X_2.
- This is particularly true in situations in which we have more than two random variables. Both will be used from here on out.

Expected values of a function of a random variable

- Definition: Let $g(X, Y)$ be a function of a bivariate random vector (X, Y).

  (a) If X and Y are discrete with joint pmf $f(x, y)$,

  $$E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

  (b) If X and Y are continuous with joint pdf $f(x, y)$,

  $$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

Examples

1. Back to the investor with two asset random variables X and Y:

$y \backslash x$	90	100	110
0	.05	.27	.18
10	.15	.33	.02

   Find $E(X+Y)$ and $E(XY)$.

2. Suppose $f(x,y) = 3x \quad$ for $0 \leq y \leq x \leq 1$.

   Find $E(X^2Y^2)$.

Special expectations

- Just like with probabilities, even if we are working with random vectors, we may only be interested in expectations for a single variable.
- Definitions:

  (a) Let (X_1, X_2) be a bivariate discrete random vector with joint pmf $f(x_1, x_2)$.

  i) If $g(X_1, X_2) = X_1$, then

  ii) If $g(X_1, X_2) = (X_1 - \mu_1)^2$, then

  iii) If $g(X_1, X_2) = \mathrm{e}^{tX_1}$, then

  Results: The mean μ_i, variance σ_i^2 and mgf $M_{X_i}(t)$ can be computed from the joint distribution (pmf / pdf) $f(x_1, x_2)$ or the marginal distribution (pmf / pdf) $f_{X_i}(x_i)$ for $i = 1, 2$.

  (b) Same results hold in the continuous case, just replace the summation with integration.

Examples

1. Continuing with the investor with two asset random variables X and Y:

$y \backslash x$	90	100	110
0	.05	.27	.18
10	.15	.33	.02

   (a) Let $g_1(X, Y) = X$. Find $E[g_1(X, Y)]$ using the joint pmf and then using the marginal pmf.

   (b) Let $g_2(X, Y) = Y$. Find $E[g_2(X, Y)]$.

   (c) Compare $E(X + Y)$ and $E(XY)$ to their "intuitive answers".

2. Suppose $f(x, y) = 1/2 \quad$ for $0 \leq x \leq y \leq 2$ and $g(X, Y) = Y$.
   Find $E[g(X, Y)]$ both ways (joint pdf and marginal pdf).

Expected value of $X + Y$ and XY

- Theorem: **Expected value of a sum of two random variables**.
  - Let (X, Y) be a bivariate random vector and function $g(X, Y) = X + Y$.

    $E(X + Y) = E(X) + E(Y)$

    Note: This result always holds (regardless of independence).
  - Proof:

    A similar proof is used for continuous random variables, again just replace $\sum$ with $\int$.
  - We could generalize this theorem to the following:

    If $g_1(X, Y)$ and $g_2(X, Y)$ are two functions and a, b and c are constants, then

    $E[ag_1(X, Y) + bg_2(X, Y) + c] = aE[g_1(X, Y)] + bE[g_2(X, Y)] + c$
  - $E(X + Y)$ is a special case with $g_1(X, Y) = X$, $g_2(X, Y) = Y$, $a = b = 1$, $c = 0$.
- Products of random variables are not so simple.

  The expected value of XY does *not* always equal the product of the expected values.
- Theorem: **Expected value of a product of two random variables**.
  - Let (X, Y) be a bivariate random vector and function $g(X, Y) = XY$.

    If X and Y are independent, $E(XY) = E(X) \cdot E(Y)$

– Notes:

This theorem may fail to hold if X and Y are not independent. There are examples of random variables X and Y which are not independent but the results of this theorem are still true.

It is still important to be able to compute $E(XY)$ directly when X and Y are not known to be independent, and because it is used in the calculation of covariance, which is also where we will see why the theorem doesn't go both ways.