Ch2 Multivariate Distribution

Jih-Chang Yu

August 9, 2023

Distribution of Two Random Variables

Definition of Sample Space of Two Random Variables

Let X and Y be two random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. The sample space of two random variables, denoted as Ω_{XY}, is defined as the set of all possible outcomes of the pair (X, Y).

Formally, $\Omega_{XY} = \{(x, y) : x \in \Omega_X, y \in \Omega_Y\}$, where Ω_X represents the sample space of X and Ω_Y represents the sample space of Y.
OR

$$\{(x, y) : x = X(c), \, y = Y(c), c \in S.\}$$

Let X and Y be two random variables with a joint probability distribution function $P_{XY}(x, y)$ defined on the sample space Ω_{XY}.

The probability of an event $A \subseteq \Omega_{XY}$ can be calculated using the joint probability function as:

$$P(A) = P_{XY}((x, y) \in A).$$

Distribution of Two Random Variables

Let X and Y be two random variables with a joint probability distribution function $P_{XY}(x, y)$ defined on the sample space Ω_{XY}.

For discrete random variables, the probability of an event $A \subseteq \Omega_{XY}$ is calculated as:

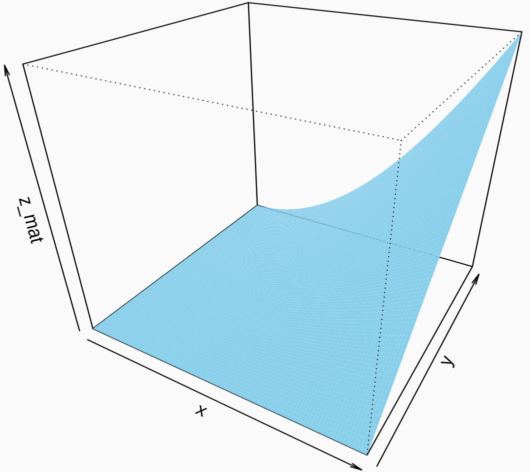
$$P(A) = \sum_{(x,y)\in A} P_{XY}(x, y)$$

where the sum is taken over all (x, y) pairs in A. For continuous random variables, the probability of an event $A \subseteq \Omega_{XY}$ is calculated using the joint probability density function $f_{XY}(x, y)$ as:

$$P(A) = \iint_{(x,y)\in A} f_{XY}(x, y)\, dx\, dy.$$

exp1

exp2



Distribution of Two Random Variables

Let X and Y be two continuous random variables with a joint cumulative distribution function (CDF) $F_{XY}(x, y)$.

The joint CDF gives the probability that both X and Y are less than or equal to their respective values x and y, and it is defined as:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

The joint probability density function (PDF) $f_{XY}(x, y)$ can be obtained by differentiating the joint CDF with respect to both variables:

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}.$$

Marginal Distribution for Discrete Random Variables

Let X and Y be two discrete random variables with a joint probability mass function (PMF) $P_{XY}(x, y)$.

The marginal distribution for X, denoted as $P_X(x)$, is obtained by summing the joint probabilities over all possible values of Y:

$$P_X(x) = \sum_y P_{XY}(x, y)$$

Similarly, the marginal distribution for Y, denoted as $P_Y(y)$, is obtained by summing the joint probabilities over all possible values of X:

$$P_Y(y) = \sum_x P_{XY}(x, y)$$

Marginal Distribution for Continuous Random Variables

Let X and Y be two continuous random variables with a joint probability density function (PDF) $f_{XY}(x, y)$.

The marginal distribution for X, denoted as $f_X(x)$, is obtained by integrating the joint PDF over the entire range of Y:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$$

Similarly, the marginal distribution for Y, denoted as $f_Y(y)$, is obtained by integrating the joint PDF over the entire range of X:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx$$

exp3

exp4

Conditional Distributions and Expectations

Conditional PMF for Discrete Random Variables

Let X and Y be two discrete random variables with a joint PMF $P_{XY}(x, y)$. The conditional PMF of X given $Y = y$, denoted as $P_{X|Y}(x|y)$, is calculated as:

$$P_{X|Y}(x|y) = \frac{P_{XY}(x, y)}{P_Y(y)}$$

where $P_Y(y)$ is the marginal probability mass function of Y.

Conditional Expectation for Discrete Random Variables

Let X and Y be two discrete random variables with a joint PMF $P_{XY}(x, y)$. The conditional expectation of X given $Y = y$, denoted as $E[X|Y = y]$, is calculated as:

$$E[X|Y = y] = \sum_{x} x \cdot P_{X|Y}(x|y)$$

where $P_{X|Y}(x|y)$ is the conditional PMF of X given $Y = y$.

Conditional Variance for Discrete Random Variables

Let X and Y be two discrete random variables with a joint PMF $P_{XY}(x, y)$. The conditional variance of X given $Y = y$, denoted as $Var(X|Y = y)$, is calculated as:

$$Var(X|Y = y) = E[(X - E[X|Y = y])^2|Y = y]$$

where $E[X|Y = y]$ is the conditional expectation of X given $Y = y$.

Conditional PDF for Continuous Random Variables

Let X and Y be two continuous random variables with a joint PDF $f_{XY}(x, y)$. The conditional PDF of X given $Y = y$, denoted as $f_{X|Y}(x|y)$, is calculated as:

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

where $f_Y(y)$ is the marginal probability density function of Y.

Conditional Expectation for Continuous Random Variables

Let X and Y be two continuous random variables with a joint PDF $f_{XY}(x, y)$. The conditional expectation of X given $Y = y$, denoted as $E[X|Y = y]$, is calculated as:

$$E[X|Y = y] = \int x \cdot f_{X|Y}(x|y)\, dx$$

where $f_{X|Y}(x|y)$ is the conditional PDF of X given $Y = y$.

Conditional Variance for Continuous Random Variables

Let X and Y be two continuous random variables with a joint PDF $f_{XY}(x, y)$. The conditional variance of X given $Y = y$, denoted as $Var(X|Y = y)$, is calculated as:

$$Var(X|Y = y) = E[(X - E[X|Y = y])^2|Y = y]$$

where $E[X|Y = y]$ is the conditional expectation of X given $Y = y$.

exp5

exp6

Correlation Coefficient

Covariance

The covariance between two random variables X and Y is a measure of the linear relationship between them. It is denoted as $Cov(X, Y)$ and defined as:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

where $E[X]$ and $E[Y]$ are the expectations (means) of X and Y, respectively.

The covariance can take positive, negative, or zero values. A positive covariance indicates a positive linear relationship, a negative covariance indicates a negative linear relationship, and a covariance of zero indicates no linear relationship between the variables.

Correlation Coefficient

The correlation coefficient between two random variables X and Y measures the strength and direction of their linear relationship. It is denoted as $Corr(X, Y)$ or ρ_{XY} and is defined as:

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

where $Cov(X, Y)$ is the covariance between X and Y, and $Var(X)$ and $Var(Y)$ are the variances of X and Y, respectively.

The correlation coefficient takes values between -1 and 1. A correlation coefficient of 1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship between the variables.

exp7

Independence of Random Variables

If X and Y are independent, it implies that there is no linear relationship between them. In other words, the knowledge of one variable provides no information about the other variable, and their joint distribution can be factorized as the product of their marginal distributions:

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$$

for continuous random variables, or

$$P_{XY}(x, y) = P_X(x) \cdot P_Y(y)$$

for discrete random variables.

Property of Expectation: Product of Two Random Variables

If X and Y are independent random variables, their joint distribution can be factorized, and the expectation simplifies to:

$$E[XY] = E[X] \cdot E[Y]$$

$$Cov(XY) = 0$$

Installation and Environment Setup

- Download and install R and RStudio.
- Configure working directory and environment variables.

Basic Syntax, Data Structures, and Control Flow

- Introduce R's basic syntax rules.
- Explain common data types: numbers, strings, vectors, lists, data frames.
- Assign values, and perform math operations.
- Explore data structures: vectors, lists, data frames.
- Control flow with if statements and loops.

Functions, Data Manipulation, Visualization

- Define and use functions in R.
- Utilize built-in functions for data manipulation.

Data Analysis Examples and Conclusion

- Practical data analysis examples: means, medians, scatter plots.
- Referencing R's official documentation and online tutorials.
- Concluding remarks on the significance of R programming skills.

Basic Syntax: Introducing R's Syntax Rules

R's syntax follows certain rules for commands and expressions:

- Commands are terminated by a newline or semicolon.
- # is used for comments.
- Objects are assigned using $< -$ or $=$.

Assign

Listing 1: R code

```
# Assigning a value of 1 to the variable x
x <- 2
# Creating a vector with elements 1, 2, and 3
x <- c(1, 2, 3)
x<-1:6
# Creating a 2x3 matrix with values 1 through 6
x <- matrix(c(1:6), 2, 3)
```

help

?function: Provides help for founction R

Listing 2: R code

```
?matrix
?sqrt
?runif
?rnorm
```

Example in R: Adding Values with the "+" Operator (1/4)

Let's demonstrate adding values using the '+' operator in R:

Listing 3: R code example

```
# Adding numeric values
value1 <- 15
value2 <- 7
sum_result <- value1 + value2
```

Example in R: Adding Values with the "+" Operator (2/4)

Adding elements of a vector:

Listing 4: R code example

```
vector1 <- c(2, 4, 6, 8, 10)
vector2 <- c(1, 3, 5, 7, 9)
vector_sum <- vector1 + vector2
```

Example in R: Adding Values with the "+" Operator (3/4)

Adding elements of a matrix:

Listing 5: R code example

```
matrix1 <- matrix(1:6, nrow = 2)
matrix2 <- matrix(7:12, nrow = 2)
matrix_sum <- matrix1 + matrix2
```

Example in R: Multiplying Value using the "*" Operator (1/4)

Let's demonstrate multiplying a value using the '*' operator in R:

Listing 6: R code example

```
# Multiplying a value
value1 <- 5
value2 <- 3
product_result <- value1 * value2
```

Example in R: Multiplying Elements of a Vector using the "*" Operator (2/4)

Let's demonstrate multiplying elements of a vector using the '*' operator in R:

Listing 7: R code example

```
# Multiplying elements of a vector
vector1 <- c(2, 4, 6, 8, 10)
vector2 <- c(1, 3, 5, 7, 9)
vector_product <- vector1 * vector2
```

Example in R: Multiplying Elements of a Matrix using the "*" Operator (3/4)

Let's demonstrate multiplying elements of a matrix using the '*' operator in R:

Listing 8: R code example

```r
# Multiplying elements of a matrix
matrix1 <- matrix(1:6, nrow = 2)
matrix2 <- matrix(7:12, nrow = 2)
matrix_product <- matrix1 * matrix2
```

Example in R: Using the Exponentiation Operator with a Value (1/3)

Let's demonstrate using the Exponentiation Operator (^) with a value in R:

Listing 9: R code example

```
# Using the Exponentiation Operator with a value
value <- 3
exponent <- 4
result <- value^exponent
```

Example in R: Using the Exponentiation Operator with a Vector (2/3)

Let's demonstrate using the Exponentiation Operator (ˆ) with a vector in R:

Listing 10: R code example

```
# Using the Exponentiation Operator with a vector
vector <- c(2, 3, 4)
exponent <- 2
result <- vector^exponent
```

Example in R: Using the Exponentiation Operator with a Matrix (3/3)

Let's demonstrate using the Exponentiation Operator (^) with a matrix in R:

Listing 11: R code example

```
# Using the Exponentiation Operator with a matrix
M <- matrix(1:4, nrow = 2)
exponent <- 3
result <- M^exponent
```