Chapter 3

Probability

The Basis of the Statistical inference

- **Key words:**

- Probability, objective Probability,

subjective Probability, equally likely

Mutually exclusive, multiplicative rule

Conditional Probability, independent events, Bayes theorem

3.1 INTRODUCTION

- The concept of probability is frequently encountered in everyday communication. **For example**, a physician may say that a patient has a 50-50 chance of surviving a certain operation.
  Another physician may say that she is 95 percent certain that a patient has a particular disease.
- Most people express probabilities in terms of percentages.
- But, it is more convenient to express probabilities as fractions. Thus, we may measure the probability of the occurrence of some event by a number between 0 and 1.
- The more likely the event, the closer the number is to one. An event that can't occur has a probability of zero, and an event that is certain to occur has a probability of one.

3.2 TWO VIEWS OF PROBABILITY OBJECTIVE AND SUBJECTIVE:

- *** Objective Probability
- ** Classical and Relative
- Some definitions:

1.Equally likely outcomes:

Are the outcomes that have the same chance of occurring.

2.Mutually exclusive:

Two events are said to be mutually exclusive if they cannot occur simultaneously such that $A \cap B = \Phi$.

- The universal Set (S): The set all possible outcomes.
- The empty set Φ : Contain no elements.
- The event ,E : is a set of outcomes in S which has a certain characteristic.
- Classical Probability : If an event can occur in N mutually exclusive and equally likely ways, and if m of these possess a triat, E, the probability of the occurrence of event E is equal to m/ N .
- For Example: in the rolling of the die , each of the six sides is equally likely to be observed . So, the probability that a 4 will be observed is equal to 1/6.

- Relative Frequency Probability:
- Def: If some posses is repeated a large number of times, n, and if some resulting event E occurs m times , the relative frequency of occurrence of E , m/n will be approximately equal to probability of E . $P(E) = m/n$.
- *** Subjective Probability :
- Probability measures the confidence that a particular individual has in the truth of a particular proposition.
- For Example : the probability that a cure for cancer will be discovered within the next 10 years.

3.3 ELEMENTARY PROPERTIES OF PROBABILITY:

- Given some process (or experiment) with n mutually exclusive events $E_1, E_2, E_3, \ldots\ldots\ldots\ldots, E_n$, then
- 1-$P(E_i)$ 0, i= 1,2,3,……n
- 2- $P(E_1) + P(E_2) + \ldots\ldots + P(E_n) = 1$
- 3- $P(E_i + E_J) = P(E_i) + P(E_J)$, E_i, E_J are mutually exclusive
-

RULES OF PROBABILITY

- 1-Addition Rule
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- 2- If A and B are mutually exclusive (disjoint) ,then
- $P(A \cap B) = 0$
- Then , addition rule is
- $P(A\ B) = P(A) + P(B)$.
- 3- Complementary Rule
- $P(A') = 1 - P(A)$
- where, A' = = complement event
- Consider example 3.4.1 Page 63

TABLE 3.4.1 IN EXAMPLE 3.4.1

Family history of Mood Disorders	Early = 18 (E)	Later >18 (L)	Total
Negative(A)	28	35	63
Bipolar Disorder(B)	19	38	57
Unipolar (C)	41	44	85
Unipolar and Bipolar(D)	53	60	113
Total	141	177	318

**ANSWER THE FOLLOWING QUESTIONS:

Suppose we pick a person at random from this sample.

1-The probability that this person will be 18-years old or younger?

2-The probability that this person has family history of mood orders Unipolar(C)?

3-The probability that this person has no family history of mood orders Unipolar( )?

4-The probability that this person is 18-years old or younger or has no family history of mood orders Negative (A)?

5-The probability that this person is more than18-years old $\overline{C}$ and has family history of mood orders Unipolar and Bipolar(D)?

CONDITIONAL PROBABILITY:

P(A|B) is the probability of A assuming that B has happened.

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$
- $P(B|A) = \frac{P(A \cap B)}{P(A)}$, $P(A) \neq 0$

EXAMPLE 3.4.2 PAGE 64

From previous example 3.4.1 Page 63 , answer

- suppose we pick a person at random and find he is 18 years or younger (E),what is the probability that this person will be one who has no family history of mood disorders (A)?
- suppose we pick a person at random and find he has family history of mood (D) what is the probability that this person will be 18 years or younger (E)?

CALCULATING A JOINT PROBABILITY :

- Example 3.4.3.Page 64
- Suppose we pick a person at random from the 318 subjects. Find the probability that he will early (E) and has no family history of mood disorders (A).

MULTIPLICATIVE RULE:

- $P(A \cap B) = P(A|B)P(B)$
- $P(A \cap B) = P(B|A)P(A)$
- Where,
- $P(A)$: marginal probability of A.
- $P(B)$: marginal probability of B.
- $P(B|A)$:The conditional probability.

EXAMPLE 3.4.4 PAGE 65

- From previous example 3.4.1 Page 63 , we wish to compute the joint probability of Early age at onset(E) and a negative family history of mood disorders(A) from a knowledge of an appropriate marginal probability and an appropriate conditional probability.
- Exercise: Example 3.4.5.Page 66
- Exercise: Example 3.4.6.Page 67

INDEPENDENT EVENTS:

- If A has no effect on B, we said that A,B are independent events.
- Then,
- 1- $P(A \cap B) = P(B)P(A)$
- 2- $P(A|B)=P(A)$
- 3- $P(B|A)=P(B)$

EXAMPLE 3.4.7 PAGE 68

- In a certain high school class consisting of 60 girls and 40 boys, it is observed that 24 girls and 16 boys wear eyeglasses . If a student is picked at random from this class ,the probability that the student wears eyeglasses , P(E), is 40/100 or 0.4 .
- What is the probability that a student picked at random wears eyeglasses given that the student is a boy?
- What is the probability of the joint occurrence of the events of wearing eye glasses and being a boy?

EXAMPLE 3.4.8 PAGE 69

- Suppose that of 1200 admission to a general hospital during a certain period of time,750 are private admissions. If we designate these as a set A, then compute P(A) , P( ).

$\bar{A}$

- Exercise: Example 3.4.9.Page 76

MARGINAL PROBABILITY:

- Definition:
- Given some variable that can be broken down into m categories designated

by $A_1, A_2, \ldots\ldots, A_i, \ldots\ldots, A_m$ and another jointly occurring variable that is broken down into n categories designated by $B_1, B_2, \ldots\ldots, B_j, \ldots\ldots, B_n$, the marginal probability of with all the categories of B . That is,

$$P(A_i) = \sum P(A_i \cap B_j),$$ for all value of j

- Example 3.4.9.Page 76
- Use data of Table 3.4.1, and rule of marginal Probabilities to calculate P(E).

EXERCISE:

- Page 76-77
- Questions :
- 3.4.1, 3.4.3,3.4.4
- H.W.
- 3.4.5 , 3.4.7

BAYE'S THEOREM
PAGES 79-83

Test Result	Yes (D)	No ($\overline{D}$)	Tota
Positive(T)	a	b	a+b
Negative ($\overline{\mathrm{T}}$)	c	d	c+d
Total	a+c	b+d	N

Definition 1

A False Positive Result

A false positive is when the test indicates a positive result (T) when the person does not have the disease $\overline{D}$

Probability of a A False Positive Result = P(T | $\overline{D}$)

Definition 2

A False Negative Result

A false negative is when a test indicates a negative result ($\overline{\mathrm{T}}$) when the person has the disease (D).

Probability of a A False Negative Result = P($\overline{\mathrm{T}}$ | D)

Test Result	Yes (D)	No ($\overline{D}$)	Tota
Positive(T)	a	b	a+b
Negative ($\overline{T}$)	c	d	c+d
Total	a+c	b+d	N

Definition.3

The sensitivity of the symptom

This is the probability of a positive result given that the subject has the disease. It is denoted by $P(T|D)$

Definition.4

The specificity of the symptom

This is the probability of negative result given that the subject does not have the disease. It is denoted by

Definition.5

The predictive value positive of the symptom

This is the probability that a subject has the disease given that the subject has a positive screening test result

It is calculated using Bayes Theorem through the following formula

$$P(D \mid T) = \frac{P(T \mid D)P(D)}{P(T \mid D)P(D) + P(T \mid \overline{D})P(\overline{D})}$$

Where P(D) is the rate of the disease which is always given.

$$P(\overline{D}) = 1 - P(D)$$

$$p(T \mid \overline{D}) = 1 - P(\overline{T} \mid \overline{D})$$

Note that: the numerator is equal to the sensitivity times rate of the disease; while the denominator is equal to the sensitivity times the rate of the disease plus 1 minus the specifity times 1 minus the rate of the disease.

Definition.6

The predictive value negative of the symptom

This is the probability that a subject does not have the disease given that the subject has a negative screening test result
It is calculated using Bayes Theorem through the following formula

$$P(\overline{D} \mid \overline{T}) = \frac{P(\overline{T} \mid \overline{D})P(\overline{D})}{P(\overline{T} \mid \overline{D})P(\overline{D}) + P(\overline{T} \mid D)P(D)}$$

where,

$$p(\overline{T} \mid D) = 1 - P(T \mid D)$$

Example 3.5.1 page 82

A medical research team wished to evaluate a proposed screening test for Alzheimer's disease. The test was given to a random sample of 450 patients with Alzheimer's disease and an independent random sample of 500 patients without symptoms of the disease. The two samples were drawn from populations of subjects who were 65 years or older. The results are as follows.

Test Result	Yes (D)	No ($\overline{D}$)	Total
Positive(T)	436	5	441
Negativ($\overline{T}$)	14	495	509
Total	450	500	950

In the context of this example

a) Probability of a A False Positive Result = P(T| $\overline{D}$)=(5/500

b) Probability of a A False Negative Result = P($\overline{\mathrm{T}}$ | D) =14/450

c) Compute the sensitivity of the symptom.

$$P(T \mid D) = \frac{436}{450} = 0.9689$$

d) Compute the specificity of the symptom.

$$P(\overline{T} \mid \overline{D}) = \frac{495}{500} = 0.99$$

e) Suppose it is known that the rate of the disease in the general population is 11.3%. What is the predictive value positive of the symptom and the predictive value negative of the symptom

The predictive value positive of the symptom is calculated as

$$P(D \mid T) = \frac{P(T \mid D)P(D)}{P(T \mid D)P(D) + P(T \mid \overline{D})P(\overline{D})}$$

$$= \frac{(0.9689)(0\ .113)}{(0.9689)(0\ .113) + (.01)(1 - 0.113)} = 0.925$$

The predictive value negative of the symptom is calculated as

$$P(\overline{D} \mid \overline{T}) = \frac{P(\overline{T} \mid \overline{D})P(\overline{D})}{P(\overline{T} \mid \overline{D})P(\overline{D}) + P(\overline{T} \mid D)P(D)}$$

$$= \frac{(0.99)(0.887)}{(0.99)(0.887) + (0.0311)(0.113)} = 0.996$$

EXERCISE:

- Page 83
- Questions :
- 3.5.1, 3.5.2
- H.W.:
- Page 87 : Q4,Q5,Q7,Q9,Q21