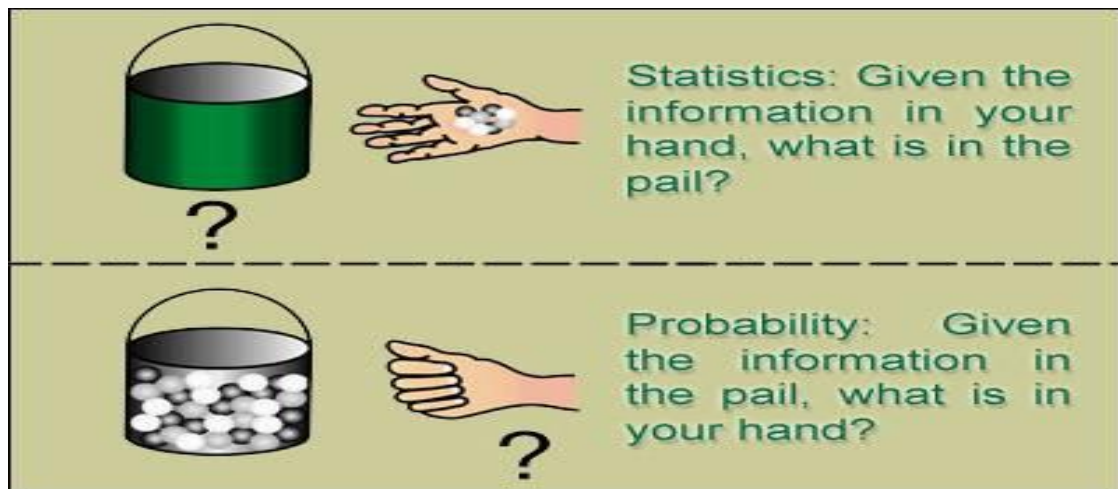


Chapter three (statistics)



Statistical methods can be used to find answers to the questions like:

- What kind and how much data need to be collected?
- How should we organize and summarize the data?
- How can we analyse the data and draw conclusions from it?
- How can we assess the strength of the conclusions and evaluate their uncertainty?

That is, statistics provides methods for

1. Design: Planning and carrying out research studies.
2. Description: Summarizing and exploring data.
3. Inference: Making predictions and generalizing about phenomena represented by the data.

Statistic Divided to

1. Data Description
2. Statistical Inference
3. Forecasting

Random variables and probability distribution

المتغيرات العشوائية والتوزيعات الاحتمالية

In this chapter, we will presented

- ✚ definition of random variables
- ✚ probability mass functions
- ✚ probability density function
- ✚ expected values, mean, and variance
- ✚ example distributions: binomial, exponential, Poisson, Gaussian

Variables A characteristic that varies from one person or thing to another is called a variable,

i.e, a variable is any characteristic that varies from one individual member of the population to another.

Examples of variables for humans are height, weight, and number of siblings, sex, marital status, and eye color. The first three of these variables yield numerical information (yield numerical measurements) and are examples of quantitative (or numerical) variables; last three yield non-numerical information (yield non-numerical measurements) and are examples of qualitative (or categorical) variables. Quantitative variables can be classified as either discrete or continuous.

Discrete variables. Some variables, such as the numbers of children in family, the numbers of car accident on the certain road on different days, or the numbers of students taking basics of statistics course are the results of counting and thus these are discrete variables.

Definition:- (Notation for parameters). Let μ and σ denote the mean and standard deviation of a variable for the population, we call μ and σ the population mean and population standard deviation.

Definition :- (Random variable). A random variable is a variable whose value is a numerical outcome of a random phenomenon.

Or a random variable is a function that associates a real number with each element in the sample space.

We usually denote random variables by capital letters near the end of the alphabet, such as X or Y . Some values of the random variable X may be more likely than others. The probability distribution of the random variable X lists the possible outcomes together with their probabilities the variable X can have.

The probability distribution

The probability distribution of a discrete random variable X assigns a probability to each possible values of the variable.

Each probability is a number between 0 and 1, and the sum of the probabilities of all possible values equals 1.

Let $x_i, i = 1, 2, \dots, k$, denote a possible outcome for the random variable X,

And let $P(X = x_i) = P(x_i) = P_i$ denote the probability of that outcome.

Then $0 \leq P(x_i) \leq 1$ and $\sum_{i=1}^k p(x_i) = 1$.

since each probability falls between 0 and 1, and since the total probability equals 1.

Definition:- (Probability distribution of a discrete random variable) Or probability mass function

A discrete random variable X has a countable number of possible values. The probability distribution of X lists the values and their probabilities:

Value of X	$x_1, x_2, x_3, \dots, x_k$
Probability	$p(x_1), p(x_2), p(x_3), \dots, p(x_k)$

The probabilities $P(x_i)$ must satisfy two requirements:

1. Every probability $P(x_i)$ is a number between 0 and 1.
2. $P(x_1) + P(x_2) + \dots + P(x_k) = 1$.

Example :- The instructor of a large class gives 15% each of 5=excellent, 20% each of 4=very good, 30% each of 3=good, 20% each of 2=satisfactory, 10% each of 1=sufficient, and 5% each of 0=fail. Choose a student at random from this class.

The student's grade is a random variable X. The value of X changes when we repeatedly choose students at random, but it is always one of 0,1,2,3,4 or 5. What is the probability distribution of X? Draw a probability histogram for X. What is the probability that the student got 4=very good or better, i.e, $P(X \geq 4)$?