Essentials of Data Science With R Software - 1

Probability Theory and Statistical Inference

Univariate Random Variables

:::

Lecture 25
Expectation of Variables

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Expectation of a Random Variable: Example 1

What does expectation means?

Suppose we toss a coin – two possible outcomes –

Head (*H*) and Tail (*T*).

Suppose we decide

- **if we get Head, we get a reward of Rs. 2 and**
- **if we get Tail, we get a reward of Rs 4.**

What do we expect to get an average reward?

Note that $P(H) = P(T) = \frac{1}{2}$

Expected average reward = Rs. $2 \times \frac{1}{2}$ + Rs. $4 \times \frac{1}{2}$ = Rs. 3

Expectation of a Random Variable: Example 2

Suppose we roll a dice and following is the scheme for award based on outcomes –

Point (x)	**1**	**2**	**3**	**4**	**5**	**6**
Reward (INR)	**1**	**2**	**3**	**4**	**5**	**6**
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Expected reward money =

$$\text{Rs. } 1 \times \frac{1}{6} + \text{Rs. } 2 \times \frac{1}{6} + \text{Rs. } 3 \times \frac{1}{6} + \text{Rs. } 4 \times \frac{1}{6} + \text{Rs. } 5 \times \frac{1}{6} + \text{Rs. } 6 \times \frac{1}{6}$$

= Rs. 3.50

Mathematical Expectation of a Continuous Random Variable:

Let *X* be a continuous random variable having the probability density function *f* (*x*).

Suppose *g*(*X*) is a real valued function of *X*.

Obviously *g*(*X*) will also be a random variable.

Then expectation of *g*(*X*) is defined as

$$\boldsymbol{E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx}$$

provided with $\int |g(x)| f(x)\, d\mathrm{x} < \infty$**.**

Mathematical Expectation of a Discrete Random Variable:

Let X be a discrete random variable having the probability mass function $P(X = x_i) = p_i$.

Suppose $g(X)$ is a real valued function of X.

Obviously $g(X)$ will also be a random variable.

Thus X takes the values $x_1, x_2, \ldots, x_k, \ldots$, with respective probabilities $p_1, p_2, \ldots, p_k, \ldots$

Then expectation of $g(X)$ exists and is defined as

$$E[g(X)] = \sum_{i=1}^{\infty} g(x_i)P(X = x_i) = \sum_{i=1}^{\infty} g(x_i)p_i$$

provided $\sum_{i=1}^{\infty} |g(x_i)|p_i < \infty$.

Special cases of Expectation of a Random Variable: Mean

- $g(X) = X$ then $E[g(X)] = E(X)$

The *expectation* of X, i.e. $E(X)$, is usually denoted by $\mu = E(X)$ and relates to the arithmetic mean of the distribution of the population. It reflects the central tendency of the population.

- If a and b are any real constants, then

$$E(a) = a \quad \text{and} \quad E[aX + b] = aE[X] + b$$

- Let $g_1, g_2, \ldots, g_r$ be r real valued functions such that $E[g_i(X)]$ exists for all $i = 1,2,\ldots,r$ then $E[\sum_{i=1}^{r} g_i(X)] = \sum_{i=1}^{r} E[g(X_i)]$

Special cases of Expectation of a Random Variable: Mean

Example 1

Consider the continuous random variable “waiting time for the train”. Suppose that a train arrives every 20 min. Therefore, the waiting time of a particular person is random and can be any time contained in the interval [0, 20].

The required probability density function is

$$f(x) = \begin{cases} \dfrac{1}{20} & \text{for } 0 \leq x \leq 20 \\ 0 & \text{otherwise.} \end{cases}$$

Special cases of Expectation of a Random Variable: Mean

Example 1

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

$$= \int_{-\infty}^{0} xf(x)dx + \int_{0}^{20} xf(x)dx + \int_{20}^{\infty} xf(x)dx$$

$$= 0 + \int_{0}^{20} x\frac{1}{20}dx + 0 = 10$$

Thus the "average" waiting time for the train is 10 min.

This means that if a person has to wait for the train every day, then the waiting time will vary randomly between 0 and 20 minutes and, on average, it will be 10 minutes.

Special cases of Expectation of a Random Variable: Mean

Example 2

Suppose we roll a dice and following is the scheme for award based on outcomes –

Point (x)	**1**	**2**	**3**	**4**	**5**	**6**
Reward (INR)	**1**	**2**	**3**	**4**	**5**	**6**
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(X) = \text{Rs. } 1 \times \frac{1}{6} + \text{Rs. } 2 \times \frac{1}{6} + \text{Rs. } 3 \times \frac{1}{6} + \text{Rs. } 4 \times \frac{1}{6} + \text{Rs. } 5 \times \frac{1}{6} + \text{Rs. } 6 \times \frac{1}{6}$$

$$= \text{Rs. } 3.50$$

Special cases of Expectation of a Random Variable: Mean

The arithmetic mean of observations $x_1,x_2,\ldots,x_n$ is defined as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

`mean(x)` **provides the value of arithmetic mean of the data in data vector** `x` **.**

Special cases of Expectation of a Random Variable: Mean

Example:

Following are the marks obtained out of maximum marks 100 by 20 participants in an examination: 42, 35, 45, 88, 74, 65, 78, 68, 39, 56, 76, 75, 62, 48, 62, 67, 76, 37, 58, 68.

```
> marks = c(32, 35, 45, 83, 74, 55, 68, 38, 35,
55, 66, 65, 42, 68, 72, 84, 67, 36, 42, 58)

> mean(marks)

[1] 56
```

R Console

```
> marks
 [1] 32 35 45 83 74 55 68 38 35 55 66 65 42 68 72 84 67 36 42 58
> mean(marks)
[1] 56
> 
```