Probability and Statistics (MA2001)
(Expectation, moments, skewness, kurtosis and moment generating function)

Lectures-12, 13, 14, 15

by

Dr. S. Kayal

Department of Mathematics
National Institute of Technology Rourkela
Rourkela - 769008, Odisha, India

Expectation or Mean

Definition

- Let X be a discrete type random variable with probability mass function f_X and support S_X. We say that the expected value of X (denoted by $E(X)$) is finite and equals

$$E(X) = \sum_{x \in S_X} x f_X(x)$$

provided $\sum_{x \in S_X} |x| f_X(x) < \infty$.

- Let X be a continuous type random variable with probability density function f_X. We say that the expected value of X (denoted by $E(X)$) is finite and equals

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

provided $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$.

Example

Let X be a random variable with probability mass function

$$f_X(x) = \begin{cases} (\frac{1}{2})^x, & x \in \{1, 2, 3, \cdots\} \\ 0, & \text{otherwise.} \end{cases}$$

Show that $E(X)$ exists.

Example

Let X be a random variable with probability mass function

$$f_X(x) = \begin{cases} \frac{6}{\pi^2 x^2}, & x \in \{-1, +1, -2, +2, -3, +3, \cdots\} \\ 0, & \text{otherwise.} \end{cases}$$

Show that $E(X)$ does not exist.

Solution

Give your attention into the class lecture!

Example

Let X be a random variable with probability density function

$$f_X(x) = \begin{cases} \frac{e^{-|x|}}{2}, & -\infty < x < \infty \\ 0, & otherwise. \end{cases}$$

Show that $E(X)$ is finite and find its value.

Example

Let X be a random variable with probability density function

$$f_X(x) = \begin{cases} \frac{1}{\pi}\frac{1}{1+x^2}, & -\infty < x < \infty \\ 0, & otherwise. \end{cases}$$

Show that $E(X)$ is not finite.

Solution

See during my lecture!

Theorem

- Let X be a random variable of discrete type with support S_X and probability mass function f_X. Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel function and let $T = h(X)$. Then,

$$E(T) = \sum_{x \in S_X} h(x) f_X(x),$$

provided it is finite.

- Let X be a random variable of (absolutely) continuous type with probability density function f_X. Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel function and let $T = h(X)$. Then,

$$E(T) = \int_{-\infty}^{\infty} h(x) f_X(x) dx,$$

provided it is finite.

Example

Let X be a random variable with probability mass function

$$f_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & x \in \{0, 1, 2, \cdots\} \\ 0, & \text{otherwise.} \end{cases}$$

Find $E(X^2)$.

Example

Let X be a random variable with probability density function

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Find $E(X^r),\ r > 0$.

Solution

Give your attention into the class lecture!

Expectation and moments

Definition

Let X be a random variable defined on some probability space.

- $\mu_1' = E(X)$, provided it is finite, is called the mean of the random variable X.
- For $r \in \{1, 2, \cdots\}$, $\mu_r' = E(X^r)$, provided it is finite, is called the rth moment of X.
- For $r \in \{1, 2, \cdots\}$, $\mu_r = E((X - \mu_1')^r)$, provided it is finite, is called the rth central moment of X.
- $\mu_2 = E((X - \mu_1')^2)$, provided it is finite, is called the variance of X. We denote $Var(X) = E((X - \mu_1')^2)$. The quantity $\sigma = \sqrt{\mu_2} = \sqrt{E((X - \mu_1')^2)}$ is called the standard deviation of X.

Theorem

Let X be a random variable.

- For real constants a and b, $E(aX + b) = aE(X) + b$, provided the involved expectations are finite.
- If $h_1, \cdots, h_m$ are Borel functions, then

$$E\left(\sum_{i=1}^{m} h_i(X)\right) = \sum_{i=1}^{m} E(h_i(X))$$

provided the involved expectations are finite.

Theorem

Let X be a random variable with finite first two moments and let $E(X) = \mu$. Then,

- $Var(X) = E(X^2) - (E(X))^2$
- $Var(X) \geq 0$. Moreover, $Var(X) = 0$ if and only if $P(X = \mu) = 1$
- $E(X^2) \geq (E(X))^2$ [Cauchy-Schwarz inequality]
- For any real constants a, b,

$$Var(aX + b) = a^2 Var(X).$$

Example

Let X be a random variable with probability density function

$$f_X(x) = \begin{cases} \frac{1}{2}, & -2 < x < -1 \\ \frac{x}{9}, & 0 < x < 3 \\ 0, & otherwise. \end{cases}$$

(i) If $Y_1 = \max(X, 0)$, find the mean and variance of Y_1.

(ii) If $Y_2 = 2X + 3e^{-\max(X,0)} + 4$, find $E(Y_2)$.

Solution

See during lecture!

Mean, median and mode

Mean

- The mean of a random variable X is given by $\mu_1' = E(X)$.

Median

- A real number m satisfying $F_X(m^-) \leq \frac{1}{2} \leq F_X(m)$, that is, $P(X < m) \leq \frac{1}{2} \leq P(X \leq m)$ is called the median of X.
- The median of probability distribution divides S_X into two equal parts each having the same probability of occurrence.
- If X is continuous, then median m is given by $F_X(m) = \frac{1}{2}$.
- For discrete case, the median may not be unique.

Mode

- The mode m_0 of a probability distribution is the value that occurs with highest probability. It is defined as $m_0 = \sup\{f_X(x) : x \in S_X\}$.

Example

Obtain mean, median and mode for X with distribution function (Ans. $3/4, (1/2)^{1/3}, 1$)

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x^3, & x \in [0,1] \\ 1, & x > 1, \end{cases}$$

Example

Consider a random variable X with probability mass function $P(X=-2) = \frac{1}{4},\ P(X=0) = \frac{1}{5},\ P(X=1) = \frac{23}{60},\ P(X=2) = \frac{1}{6}$. Obtain median of X. (Ans. 1)

Solution

See it during the lecture.

What are the importance of mean, median and mode?

Mean

- The mean or average, is calculated by summing up all the values in a data set and then dividing by the number of values.
- It gives a sense of the "typical" value in the data set.
- The mean is sensitive to outliers, which are extreme values that can significantly affect its value.
- The importance of the mean lies in its ability to provide a general sense of the central value around which the data is distributed.

Uses

- Analyzing the average package in a placement to understand the overall performance of students.
- The importance of the mean lies in its ability to provide a general sense of the central value around which the data is distributed.

Median

- The median is the middle value in a sorted data set.
- It is a measure that is less affected by outliers compared to the mean.
- If the data set has an odd number of values, the median is the middle value. If the data set has an even number of values, the median is the average of the two middle values.

Uses

- Incomes within a population, as it provides the middle value that separates the higher and lower earners.
- Housing prices, as it's less sensitive to extremely expensive or cheap properties.

Mode

- The mode is the value that appears most frequently in a data set.
- It can be a single value or a set of values if multiple values have the same highest frequency.
- The mode can help identify the most common value(s) in the data set and is particularly useful for categorical data.

Uses

- Identifying the most common eye color in a group of individuals.
- Finding the most frequent score on a test to understand the most common level of performance.

- Insurance analysts frequently compute the mean age of the people they insure in order to determine the average age of their clients.
- Actuaries frequently compute the median amount spent on healthcare by individuals each year in order to determine how much insurance they need to be able to give.
- Actuaries also calculate their clients' mode (the most common age) so they may see which age group is most likely to utilize their insurance.

Measures of skewness and kurtosis

Skewness

- Skewness of a probability distribution is a measure of asymmetry (or lack of symmetry).
- A measure of skewness of the probability distribution of X is defined as
$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}.$$
- For symmetric distribution, $\beta_1 = 0$.
- $\beta_1 > 0$ indicates that the data is positively skewed and $\beta_1 < 0$ indicates that the data is negatively skewed

Measures of skewness and kurtosis (cont...)

Kurtosis

- Kurtosis of a probability distribution of X is a measure of peakedness and thickness of tail of probability density function of X relative to the peakedness and thickness of tails of the density function of normal distribution.
- A distrbution is said to have higher (lower) kurtosis than the normal distribution if its density function in comparison with the density function of a normal distribution, has sharper (rounded) peak and longer, flatter (shorter, thinner) tails.

Measures of skewness and kurtosis (cont...)

Kurtosis (cont..)

- A measure of kurtosis of the probability distribution of X is defined by
$$\gamma_1 = \frac{\mu_4}{\mu_2^2}.$$
- For normal distribution with mean μ and variance σ^2, $\gamma_1 = 3$. The quantity
$$\gamma_2 = \gamma_1 - 3$$
is called the excess kurtosis of the distribution of X. The distribution with zero excess kurtosis is called mesokurtic. A distribution with positive (negative) excess kurtosis is called leptokurtic (platykurtic).

Skewness

Example: Income Distribution

- Description: In many countries, income distribution is often positively skewed (right-skewed). A small number of individuals earn significantly more than the majority, which pulls the tail of the distribution to the right.
- Implication: This means that while most people earn below the average income, a few high earners increase the average, reflecting a disparity in wealth.

Example: age of deaths

- Description: The distribution of the age of deaths in most populations is left-skewed. .
- Implication: This means that most people live to be between 70 and 80 years old, with fewer and fewer living less than this age

Kurtosis

Example: Exam Scores

- Description: Consider a scenario where a large number of students take a standardized test. If the scores are tightly clustered around the mean with very few low or high scores, the distribution may exhibit high kurtosis (leptokurtic). Conversely, if the scores are more evenly spread with many students scoring very low or very high, the distribution may have low kurtosis (platykurtic).
- Implication: High kurtosis indicates that there are more outliers (students scoring very high or very low), while low kurtosis suggests that the distribution of scores is more uniform with fewer extreme values.

Exercises

(i) Obtain the skewness and kurtosis of X, with probability mass function

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \ \lambda > 0, \ x = 0, 1, 2, \cdots .$$

(ii) Obtain the skewness and kurtosis of X, with probability density function

$$f_X(x) = \lambda e^{-\lambda x} \ \ x > 0, \ \lambda > 0.$$

Sol.

See the lecture!!

Moment generating function

Definition

Let X be a random variable and let
$A = \{t \in \mathbb{R} : E(|e^{tX}|) = E(e^{tX}) \text{ is finite}\}$. Define $M_X : A \to \mathbb{R}$ by

$$M_X(t) = E(e^{tX}),\ t \in \mathbb{R}$$

- We call the function $M_X(.)$ the moment generating function of X.
- We say that the moment generating function exists if there exists a positive real number a such that $(-a, a) \subset A$, that is, if $M_X(t) = E(e^{tX})$ is finite in an interval containing 0.

Theorem

Let X be a random variable with moment generating function $M_X(.)$. Then, for each $r \in \{1, 2, \cdots\}$ $\mu'_r = E(X^r) = M_X^{(r)}(0)$, where $M_X^r(.)$ is the rth derivative of $M_X(t)$.

Examples

(i) Obtain the momet generating function of X, with probability mass function

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!},\ \lambda > 0,\ x = 0, 1, 2, \cdots .$$

(ii) Obtain the moment generating function of X, with probability density function

$$f_X(x) = \lambda e^{-\lambda x} \ \ x > 0,\ \lambda > 0.$$

Sol.

See the lecture!!

Example

Let X be a discrete type random variable with moment generating function

$$M_X(t) = \frac{1}{4}e^{t} + \frac{1}{2}e^{3t} + \frac{1}{4}e^{7t}, \quad t \in \mathbb{R}.$$

Obtain the distribution function of X.

Sol.

See the lecture!!

References

- *Advanced Engineering Mathematics*, Tenth Edition, by Erwin Kreyszig, John Wiley & Sons, Inc.
- *Higher Engineering Mathematics*, Sixth Edition, by B V Ramana, Tata McGraw-Hill Publisjing Company Limited.
- *An Introduction to Probability and Statistics*, Second Edition, by V. K. Rohatgi and A. K. Md. E. Saleh, Wiley.
- *Introduction to Probability and Statistics for Engineers and Scientists* by S.M. Ross, Academic Press.

Thank You