PROBABILIY AND STATISTICS I

LECTURE THIRTEEN

Introduction to moment generating functions

Lecturer: Dr. Emily Roche

INTRODUCTION

This lecture will focus on the introduction to moments and moment generating functions.

Intended learning outcomes

At the end of this lecture, you will be able to define moment generating functions (mgf), find the mgf of random variables and use mgf to find the mean and variance of a random variable.

References

These lecture notes should be supplemented with relevant topics from the book listed in the Bibliography at the end of the lecture.

Moments and Moment generating functions (mgf)

In addition to the expectation and variance of a random variable, we can compute expectations of higher powers of a random variable with a given distribution.

As we shall partly see, these expectations called moments are useful in determining various characteristics of the corresponding probability distribution. Moments play a central role in both theoretical and applied statistics.

Definition

If X is a random variable, the r^{th} moment of X usually denoted by μ_r' is defined as

$$\mu_r' = E[X^r] \qquad r = 1, 2, 3, \ldots$$

if the expectation exists.

Note

1. If $r = 1$ $\mu_1' = E[X] =$ mean of X
2. $r = 1, 2, 3, \ldots$

These are higher order moments and are being computed about zero, the origin, and are therefore usually referred to as raw moments. This then means that moments can also be computed about another value for example the mean of a probability distribution. In this case we now have the central moments

Definition

If X is a random variable, the r^{th} moment about a value a is defined as

$$\mu_r = E[(X - a)^r]$$

Where $a = E(X) = \bar{x}$, then $\mu_1 = E[(X - \bar{X})] = E(X) - (\bar{X}) = 0$

$$\mu_2 = E[(X - \bar{X})^2] = Var(X)$$

Example

If we have a discrete random variable with probability distribution $f(x) = P[X = x]$. The moments are worked out as follows:

$$\mu_1' = E[X]$$

$$= \sum_{all\ x} xPr[X = x]$$

and the first central moment

$$\mu_1 = E[(X - \bar{X})]$$

$$= \sum_{all\ x} (X - \bar{X}) \Pr[X = x]$$

For $N -$ observations $= \frac{1}{N}\sum(X - \bar{X}) = \frac{1}{N}\sum X - \frac{N\bar{X}}{N} = \bar{X} - \bar{X} = 0$

The second and third raw moments are:

$$\mu_2' = E[X^2] = \sum_{all\ x} x^2 \Pr[X = x]$$

$$\mu_3' = E[X^3] = \sum_{all\ x} x^3 \Pr[X = x]$$

Relationship between the raw and central moments:

The second central moment is given as

$$\mu_2 = E[(X - \bar{X})^2] \qquad \text{where } \bar{X} = E(X) = \mu_1'$$

$$= E[(X - \mu_1')^2]$$

But $(X - \mu_1')^2 = X^2 - 2\mu_1' X + {\mu_1'}^2$

$$\mu_2 = E\left[X^2 - 2\mu_1' X + {\mu_1'}^2\right]$$

$$= E[X^2] - 2\mu_1' E[X] + {\mu_1'}^2$$

$$= \mu_2' - 2{\mu_1'}^2 + {\mu_1'}^2$$

$$= \mu_2' - {\mu_1'}^2$$

Indeed the 2^{nd} central moment is called the variance.

For $\mu_3 = E[(X - \bar{X})^3] = E[(X - \mu_1')^3]$

But $(X - \mu_1')^3 = X^3 - 3\mu_1' X^2 + 3{\mu_1'}^2 X - {\mu_1'}^3$

$$\mu_3 = E\left[X^3 - 3\mu_1' X^2 + 3{\mu_1'}^2 X - {\mu_1'}^3\right]$$

$$= E[X^3] - 3\mu_1' E[X^2] + 3{\mu_1'}^2 E[X] - {\mu_1'}^3$$

$$= \mu_3' - 3\mu_2'\mu_1' + 3\mu_1^3 - {\mu_1'}^3$$

$$= \mu_3' - 3\mu_2'\mu_1' + 2{\mu_1'}^3$$

Similarly show that

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'{\mu_1'}^3 - 3{\mu_1'}^4$$

As seen earlier, the probability distribution or function of a random variable usually involves one or two variables that is the mean and variance.

The third and fourth moments are extremely useful for determining the shape of a frequency distribution.

Definition

The third central moment μ_3 is sometimes called a measure of symmetry or skewness, such that:

- $\mu_3 = 0 -$ represents symmetrical distributions
- $\mu_3 > 0 -$ represents distributions skewed to the right
- $\mu_3 < 0 -$ represent distributions skewed to the left

However, knowledge of the third central moment gives no clue as to the exact shape of the distribution. Instead, we define the coefficient of skewness $\alpha_3 = \frac{\mu_3}{\sigma^3}$. This measure is unit less and will always give a zero, negative or positive value.

Definition

The fourth central moment is used as a measure of kurtosis or peakedness. It gives the degree of flatness of a frequency distribution. The coefficient of kurtosis is defined as the ratio $\alpha_4 = \frac{\mu_4}{\sigma^4}$. If α_4 is positive, then the distribution is more peaked and a flatter distribution if it is negative.

Factorial moments

If X is a random variable, the r^{th} factorial moment of X is defined as:

$$\mu_r = E[X(X-1)(X-2)(X-3)\ldots(X-r+1)]$$

for some discrete random variables. The factorial moments are usually easier to calculate than the raw moments.

Example

If we throw a die, the possible outcomes are $1, 2, 3, 4, 5, 6$ which are at unit intervals. Find the 1^{st} , 2^{nd} and 3^{rd} factorial moments.

$$\mu_1 = E[X(X-1)(X-2)(X-3)\dots(X-1+1)]$$

$$= E[X]$$

Therefore, the first factorial moment is always the mean $= \sum x\,P(x) = \frac{21}{6}$

$$\mu_2 = E[X(X-1)(X-2)(X-3)\dots(X-2+1)]$$

$$= E[X(X-1)] = \sum x\,(x-1)P(x) = \frac{70}{6}$$

$$\mu_3 = E[X(X-1)(X-2)(X-3)\dots(X-3+1)]$$

$$= E[X(X-1)(X-2) = \sum x\,(x-1)(x-2)P(x) = \frac{210}{6}$$

Moment generating functions (mgf)

Compiling all moments for a probability density function is normally a tedious task. It is therefore desirable to have a function from which all the moments can be derived at will. Such a function is called a moment generating function.

When it exists for a particular pdf, it is unique. This property enables one to determine the probability distribution given the mgf.

Definition

Let X be a random variable. The moment generating function

$$M_{(x)}(t) = E[e^{tx}] \quad \text{where} -h < t < h \qquad h > 0$$
$$t > 0$$
t is a small positive number

Case 1

If X is discrete with probability distribution $f(x)$, then

$$M_{(x)}(t) = E[e^{tx}] = \sum_{all\ x} e^{tx} f(x)$$

Case 2

If X is continuous,

$$M_{(x)}(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

The $M_{(x)}(t)$ exists if the summation in case 1 is finite and the integral in case 2 is finite.

Example

Let X be a continuous random variable with a pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & elsewhere \end{cases}$$

Find the mgf, $M_{(x)}(t)$ of X if it exists.

Solution

By definition

$$M_{(x)}(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{0}^{\infty} e^{tx} \cdot \lambda e^{-\lambda x} dx = \int_{0}^{\infty} \lambda e^{(t-\lambda)x}\, dx$$

$$= \int_{0}^{\infty} \lambda e^{-(\lambda-t)x}\, dx = \frac{-\lambda}{\lambda - t} e^{-(\lambda-t)x}\Big|_{0}^{\infty} = \frac{-\lambda}{\lambda - t}[e^{-\infty} - e^{0}] = \frac{-\lambda}{\lambda - t}\left[\frac{1}{e^{\infty}} - 1\right]$$

$$= \frac{-\lambda}{\lambda - t} \cdot -1 = \frac{\lambda}{\lambda - t}$$

Thus the mgf exists.

Derivation of moments from mgf.

Let X be a discrete random variable (the results can also be derived for a continuous random variable in an analogous manner).

$$M_{(x)}(t) = E[e^{tx}] = \sum_{all\ x} e^{tx} f(x)$$

By the series expansion e^{tx} can be expressed as the infinite series $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ therefore,

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots$$

$$M_{(x)}(t) = E[e^{tx}] = \sum_{all\ x} \left[1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots\right] f(x)$$

$$= \sum_x f(x) + \sum_x tx\, f(x) + \sum_x \frac{(tx)^2}{2!} f(x) + \sum_x \frac{(tx)^3}{3!} f(x) + \cdots$$

Differentiating with respect to t,

$$M'_{(x)}(t) = \frac{d}{dt} M_{(x)}(t) = 0 + \sum_x x\, f(x) + \sum_x tx^2\, f(x) + \sum_x \frac{t^2 x^3}{2!} f(x) + \cdots$$

Setting $t = 0$

$$M'_{(x)}(0) = \sum_x x\, f(x) = E[X] = \mu'_1$$

Getting the second derivative of $M_{(x)}(t)$ with respect to t again

$$M''_{(x)}(t) = \frac{d}{dt} M'_{(x)}(t) = \sum_x x^2\, f(x) + \sum_x tx^3\, f(x) + \sum_x \frac{t^2 x^4}{2!} f(x) + \cdots$$

Setting $t = 0$

$$M''_{(x)}(0) = \sum_x x^2\, f(x) = E[X^2] = \mu'_2$$

This then summarizes to:

$$\mu'_r = E[X^r] = M^r_{(x)}(t)\Big|_{t=0}$$

The r^{th} raw moment is derived by differentiating the mgf r times with respect to t and setting $t = 0$.

Example

Derive the expected value of X and the variance of X from the mgf in the previous example. Verify that these results are valid by computing the said quantities directly from definition.

Solution

The mgf from the previous example was found to be

$$M_{(x)}(t) = \frac{\lambda}{\lambda - t}$$

We need

$$E[X] = M'_{(x)}(t)\Big|_{t=0}$$

$$\frac{d}{dt} M_{(x)}(t) = \frac{d}{dt}\left[\frac{\lambda}{\lambda - t}\right] = \frac{d}{dt}[\lambda(\lambda - t)^{-1}] = (-\lambda)(-1)(\lambda - t)^{-2}$$

$$= \frac{\lambda}{(\lambda - t)^2}$$

When $t = 0$,

$$\frac{d}{dt} M_{(x)}(0) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} = \text{Mean}$$

$$Var\ (X) = E[X^2] - [E(X)]^2 = E[X^2] - \frac{1}{\lambda^2}$$

But

$$E[X^2] = \frac{d}{dt} M'_{(x)}(t) = \left. \frac{2\lambda}{(\lambda - t)^3} \right|_{t=0} = \frac{2\lambda}{\lambda^3} = \frac{2}{\lambda^2}$$

$$Var\ (X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Verify this from definition of Expectation and Variance as $\boldsymbol{E(X) = \int_0^\infty x\lambda e^{-\lambda x}\, dx}$

Bibliography

Gupta, SP (Dr.), (2014). *Statistical methods* (43rd Ed.). Sultan Chand & Sons.

S. C. Gupta and V. K. Kapoor, (2020). *Fundamentals of mathematical Statistics* (12th Ed). Sultan Chand & Sons.