Introduction to Probability
Chapter 4: Expectation

Dr Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Expectation
2. Moments
3. Moment Generating Function (MGF)
4. Chebyshev's Inequality

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Expectation

Let X be discrete type random variable with probability distribution $(x, p(x)),\ x = x_1, x_2, \ldots$. If

$$\sum_{x=x_1}^{\infty} |x| p(x) < \infty$$

then expected value of X exist. And the expected value of X or mean of X is defined as

$$\mu = E(X) = \sum_{x=x_1}^{\infty} x p(x).$$

The varinace of X is

$$\sigma^2 = E(X-\mu)^2 = \sum_{x=x_1}^{\infty} (x-\mu)^2 p(x) = E(X^2) - (E(X))^2.$$

Expectation

Let X be continuous type random variable with PDF $f(x)$. If

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

then expected value of X exist. And the expected value of X or mean of X is defined as

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

The varinace of X is

$$\sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = E(X^2) - (E(X))^2.$$

Example

A random variable X has the following PMF

x	-2	-1	0	1	2	3
p(x)	0.1	k	0.2	k	0.3	k

Find (i) the value of k. (ii) $E(X)$ and $Var(X)$.
(i) Since $0.1 + k + 0.2 + k + 0.3 + k = 1 \Rightarrow k = 0.133$.
(ii) $E(X) = -2 \times 0.1 - 1 \times k + 0 + 1 \times k + 2 \times 0.3 + 3 \times k = 0.799$
$E(X^2) = 4 \times 0.1 + 1 \times k + 0 + 1 \times k + 4 \times 0.3 + 9 \times k = 3.063$
$Var(X) = E(X^2) - (E(X))^2 = 2.42$.

Example

Let X be the time to failure (in days) of an electronic device has the following PDF

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

Then the mean life or average life of electronic device is

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{0}^{\infty} x \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{(\lambda)^2} \\ &= \frac{1}{\lambda}. \end{aligned}$$

Moments

- Let X be discrete type random variable with probability distribution $(x, p(x)),\ x = x_1, x_2, \ldots$, then the origin moment or rth moment about origin is $\mu'_r = E(X^r) = \sum_{x=x_1}^{\infty} x^r p(x),\ r = 1, 2, \ldots$. The central moment or rth moment about mean is $\mu_r = E(X-\mu)^r = \sum_{x=x_1}^{\infty} (x-\mu)^r p(x),\ r = 1, 2, \ldots$.

- Let X be continuous type random variable with PDF $f(x)$,then the origin moment or rth moment about origin is $\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx,\ r = 1, 2, \ldots$. The central moment or rth moment about mean is $\mu_r = E(X-\mu)^r = \int_{-\infty}^{\infty} (x-\mu)^r f(x) dx,\ r = 1, 2, \ldots$.

Relationship between central moment and origin moment

$$
\begin{aligned}
\mu_r &= E(X-\mu)^r \\
&= \sum_{j=0}^{r} \binom{r}{j} (-1)^j \mu^j \mu'_{r-j}
\end{aligned}
$$

Moment Generating Function (MGF)

Let X be random variable, then the Moment Generating Function (MGF) is defined as $M_X(t) = E(e^{tX})$. Note that

$$\begin{aligned} M_X(t) &= E\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right) \\ &= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \cdots \end{aligned}$$

Hence $\frac{d^r}{dt^r} M_X(t)|_{t=0} = E(X^r) = \mu'_r$.

Example

Example

The random variable X has PMF
$P(X = x) = \binom{n}{x} p^x \; q^{n-x}, \;\; x = 0, 1, \ldots, n, \;\; p + q = 1, \;\; 0 < p < 1.$ MGF is

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x \; q^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x \; q^{n-x} \\ &= \left(q + pe^t\right)^n \end{aligned}$$

Here $\mu = E(X) = \frac{d}{dt} M_X(t)|_{t=0} = np; \;\; \mu_2' = E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = np\,(1 + (n-1)p)$ and $\mu_2 = Var(X) = \mu_2' - \mu^2 = npq.$

Example

Let X be the time to failure (in days) of an electronic device has the following PDF

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

MGF is

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} \\ &= \lambda \int_0^\infty e^{-(\lambda - t)x} \\ &= \left(1 - \frac{t}{\lambda}\right)^{-1}, \end{aligned}$$

for $t < \lambda$. Here $\mu = E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \frac{1}{\lambda}$; $\mu_2' = E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \frac{2}{\lambda^2}$ and $\mu_2 = Var(X) = \mu_2' - \mu^2 = \frac{1}{\lambda^2}$.

Example

Example

Let X be a discrete random variable with MGF $M_X(t) = \frac{e^{2t}}{6} + \frac{e^{3t}}{2} + \frac{e^{4t}}{3}$. Determine Var($X$). Note that

$$E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \left(\frac{2e^{2t}}{6} + \frac{3e^{3t}}{2} + \frac{4e^{4t}}{3} \right) |_{t=0} = \frac{19}{6}$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \left(\frac{4e^{2t}}{6} + \frac{9e^{3t}}{2} + \frac{16e^{4t}}{3} \right) |_{t=0} = \frac{63}{6}$$

Hence

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= \frac{17}{36}. \end{aligned}$$

Let X be a random variable, then for $a, b \in \mathbb{R}$,

1. $E(aX + b) = aE(X) + b$
2. $Var(aX + b) = a^2\ Var(X)$

Example

Let random variable X has mean μ and variance σ^2. Then the random variable $Z = \frac{X-\mu}{\sigma}$ has the mean

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{E(X)-\mu}{\sigma} = 0,$$

and variance

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X) = 1.$$

Here $Z = \frac{X-\mu}{\sigma}$ is called the standarized random variable.

Chebyshev's Inequality

Let X be a general random variable with mean μ and variance σ^2. Then, for any $k > 0$,

$$P\left(|X-\mu| \geq k\right) \leq \frac{\sigma^2}{k^2}.$$

or, equivalently,

$$P\left(|X-\mu| \leq k\right) \geq 1-\frac{\sigma^2}{k^2}.$$

Example

Suppose we have sampled the weights of Ponies in the local animal shelter and found that our sample has mean 40 and variance 9 pounds. Using Chebyshev's inequality, find the lower bound of the probability that the weight of Ponies is between 34 pounds to 46 pounds.

Solution: Let X be the weight of Ponies. Here mean $\mu = 40$ and variance $\sigma^2 = 9$. Then, using Chebyshev's Inequality, we have

$$\begin{aligned} P(34 < X < 46) &= P(|X - 40| < 6) \\ &\geq 1 - \frac{\sigma^2}{6^2} = 1 - \frac{9}{36} = \frac{27}{36}. \end{aligned}$$

Example

Let X be a random variable with mean μ and variance σ^2, then using Chebyshev's inequality following values may easily be observed

$$P(|X-\mu| \leq 2\sigma) \geq 1 - \frac{1}{2^2} = 0.75$$

,i.e., 75% chances is that the observed value of X lies between two standard deviations of the mean.

$$P(|X-\mu| \leq 3\sigma) \geq 1 - \frac{1}{3^2} = 0.89$$

,i.e., 89% chances is that the observed value of X lies between three standard deviations of the mean.

$$P(|X-\mu| \leq 4\sigma) \geq 1 - \frac{1}{4^2} = 0.94$$

,i.e., 94% chances is that the observed value of X lies between four standard deviations of the mean.

Summary

This chapter introduced the concept of expectation. The moments were introduced, the mean and variance are particular cases of that. The moment generating function (MGF) was introduced and examples are presented for finding mean and variance using MGF. In last the Chebyshev's inequality was presented.