Lecture 11

Property of MGF

If

$$M_X(t) = M_Y(t),$$

then X and Y are identically distributed.

Basic Properties

1. $M_X(0) = 1$
2. For a linear transformation $Y = aX + b$:

$$M_{aX+b}(t) = E[e^{t(aX+b)}] = e^{bt}\, M_X(at).$$

Characteristic Function (CF)

The **characteristic function** of a random variable X is defined as

$$\varphi_X(t) = E[e^{itX}] = M_X(it) \quad \text{(if MGF exists)}.$$

- Characteristic function always exists.
- $|\varphi_X(t)| \le 1$.

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

From inversion,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t)\, dt$$

For $X \sim \text{Exp}(1)$,

$$M_X(t) = \frac{1}{1-t}, \quad t < 1.$$

Thus,

$$\varphi_X(t) = M_X(it) = \frac{1}{1-it}.$$

This CF can be inverted to recover the pdf of the exponential random variable.

We had earlier:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{1-it} e^{-ixt}\, dt$$

Moment Inequalities

Result 1 (Markov's Inequality)

Let X be a continuous random variable with pdf $f_X(x)$. Let $h(x)$ be a non-negative function. Then,

$$P(h(X) \geq \varepsilon) \leq \frac{E[h(X)]}{\varepsilon}, \quad \varepsilon > 0.$$

Example:

Suppose the salary of people is a random variable X, but the exact distribution is unknown. If the average salary is given as $E[X]$, then

$$P(\text{Salary} \geq 50,000) \leq \frac{E[X]}{50,000}.$$

Proof

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

Split into two parts:

$$E[h(X)] = \int_{h(x)\geq\varepsilon} h(x) f(x)\, dx + \int_{h(x)<\varepsilon} h(x) f(x)\, dx$$

Note that the second integral is always non-negative, and for the first integral, since $h(x) \geq \varepsilon$,

$$\int_{h(x)\geq\varepsilon} h(x) f(x)\, dx \geq \varepsilon\, P(h(X) \geq \varepsilon).$$

Thus,

$$E[h(X)] \geq \varepsilon\, P(h(X) \geq \varepsilon).$$

$$\implies P(h(X) \geq \varepsilon) \leq \frac{E[h(X)]}{\varepsilon}.$$

This proves **Markov's Inequality**.

We know:

$$E[h(X)] \geq \int_{h(x)\geq\varepsilon} h(x) f(x)\, dx$$

Since $h(x) \geq \varepsilon$ on this domain,

$$E[h(X)] \geq \varepsilon \int_{h(x)\geq\varepsilon} f(x)\, dx$$

$$\implies E[h(X)] \geq \varepsilon\, P(h(X) \geq \varepsilon).$$

Thus,

$$P(h(X) \geq \varepsilon) \leq \frac{E[h(X)]}{\varepsilon}.$$

Equivalently,

$$P(h(X) < \varepsilon) \geq 1 - \frac{E[h(X)]}{\varepsilon}.$$

Application

Markov's Inequality provides a fundamental tool for bounding probabilities. It is the basis for deriving further inequalities, such as **Chebyshev's Inequality**.

Markov Inequality:

If we take

$$h(X) = |X|^r, \quad r > 0,$$

then for any $\varepsilon > 0$:

$$P(|X|^r \geq \varepsilon^r) \leq \frac{E[|X|^r]}{\varepsilon^r}.$$

Equivalently,

$$P(|X| \geq \varepsilon) \leq \frac{E[|X|^r]}{\varepsilon^r}.$$

This is the **general form of Markov's Inequality**.

We start with the general inequality:

$$\Pr\big(|X| \geq \varepsilon\big) \leq \frac{E\big(|X|^r\big)}{\varepsilon^r}, \quad \text{for } r > 0.$$

This can also be written as:

$$\Pr\big(|X| \geq \varepsilon\big) \leq \frac{E\left(|X|^r\right)}{\varepsilon^r}.$$

Remark: This shows that if we know any moment of X, then we can bound its tail probability. It is not necessary to only use the first moment.

Chebyshev's Inequality

Suppose X is a random variable with expectation $E(X) = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. We choose the function $h(X) = (X - \mu)^2$. Then, for any $k > 0$,

$$\Pr\big((X-\mu)^2 \geq k^2\sigma^2\big) \leq \frac{E\big((X-\mu)^2\big)}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}.$$

Hence,

$$\Pr\big(|X-\mu| \geq k\sigma\big) \leq \frac{1}{k^2}.$$

Equivalently,

$$\Pr\big(|X-\mu| < k\sigma\big) \geq 1 - \frac{1}{k^2}.$$

For $k = 2$,

$$\Pr\big(|X-\mu| < 2\sigma\big) \geq 1 - \frac{1}{4} = \frac{3}{4}.$$