LECTURE - 10

Agenda

1. Properties of Expectation and Variance
2. Tchebysheff's theorem

PROPERTIES OF EXPECTATION AND VARIANCE

Note that the variance $V(X)$ of a random variable is the average squared distance between the values of X and the expected value μ.

DEFINITION: The standard deviation of a random variable X is the square root of the variance and is given by

$$SD(X) = \sqrt{E(X-\mu)^2}$$

The standard deviation is also a measure of the variability of a random variable, but it maintains the original units of measure. It can be thought of as the size of a typical deviation between an observed outcome and the expected value.

Example: If W is the winnings in the game discussed in the previous lecture, remember

$$P(W=-10) = \frac{1}{12}, \quad P(W=-4) = \frac{1}{3}, \quad P(W=4) = \frac{5}{12},$$

$$P(W=10) = \frac{1}{6}.$$

Then,

$$V(W) = E[W^2] - (E[W])^2 \quad (\text{Why?})$$

$$= 37 - \left(\frac{7}{6}\right)^2$$

$$= 37 - \frac{49}{36}$$

$$= \frac{1283}{36}.$$

$$= 35.639$$

$$SD(W) = \sqrt{V(W)} = 5.970$$

$\boxed{(i)}$ $E(aX+b) = aE(X) + b$

Proof: $E(aX+b) = \sum_{x \in \mathcal{X}} (ax+b)\, p_X(x)$

$$= \sum_{x \in \mathcal{X}} [ax\, p_X(x) + b p_X(x)]$$

$$= \sum_{x \in \mathcal{X}} ax\, p_X(x) + \sum_{x \in \mathcal{X}} b\, p_X(x)$$

$$= aE(X) + b.$$

$\boxed{\text{(ii)}}$ $V(aX+b) = a^2V(X)$

Proof:

$$V(aX+b) = E[(aX+b) - E(aX+b)]^2$$

$$= E[(aX+b) - aE(X) - b]^2$$

$$= E[a(X - E(X))]^2$$

$$= a^2 E[(X - E(x))^2]$$

$$= a^2 V(X)$$

$\boxed{\text{(iii)}}$ $V(X) = E(X^2) - (E(X))^2$

$$V(X) = E(X-\mu)^2$$

$$= E(X^2 - 2X\mu + \mu^2)$$

$$= E(X^2) - 2\mu E(X) + E(\mu^2)$$

$$= E(X^2) - 2\mu^2 + \mu^2$$

$$= E(X^2) - \mu^2$$

RESULT: Let X be a random variable with mean μ and variance σ^2. Then for any positive k,

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

Proof: Let $V(X) = \sigma^2$ and $E(X) = \mu$

$$V(X) = \sum_{x \in \mathcal{X}} (x-\mu)^2 p_X(x)$$

$$= \sum_{x \in \mathcal{X} : |x-\mu| \geq k\sigma} (x-\mu)^2 p_X(x) + \sum_{x \in \mathcal{X} : |x-\mu| < k\sigma} (x-\mu)^2 p_X(x)$$

$$\geq \sum_{x \in \mathcal{X} : |x-\mu| \geq k\sigma} (x-\mu)^2 p_X(x)$$

$$\geq k^2 \sigma^2 \sum_{x \in \mathcal{X} : |x-\mu| \geq k\sigma} p_X(x)$$

$$= k^2 \sigma^2 P(|X-\mu| \geq k\sigma)$$

Hence, $\sigma^2 \geq k^2 \sigma^2 P(|X-\mu| \geq k\sigma)$

Hence, $\dfrac{1}{k^2} \geq P(|X-\mu| \geq k\sigma)$

This gives us the required identity as

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

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

If we choose $k = 3.16$, $1 - \frac{1}{k^2} = 90\%$.

Hence, if a random variable X has $E[X] = \mu$

and $V(X) = \sigma$, then

$$P(|X - \mu| < 3.16\sigma) \geq 90\%.$$

$$\Rightarrow P(\mu - 3.16\sigma < X < \mu + 3.16\sigma) \geq 90\%$$

This gives us a general bound, but this bound can be quite crude.