MA1061 17/11/21

Chebyshevs' Inequality

Corollary (Chebyshev's inequality)

For any random variable X and $k > 0$

$$P\{|X - E[X]| \geq k\sigma\} \leq \frac{1}{k^2}$$ where σ is the standard deviation of X.

Remark

The last one shows that deviations of X from its mean significantly greater than s.d (X) are rare.

Proof

For any Y and $c > 0$, by Markov's inequality we have

$$P\{|Y| \geq c\} = P\{Y^2 \geq c^2\} \leq \frac{E[Y^2]}{c^2}$$

Let $Y = X - E[X]$ leads to

$$P\{|X - E[X]| \geq c\} \leq \frac{E[(X - E[X])^2]}{c^2} = \frac{var(X)}{c^2}$$

Now let $c = k\sigma$ and we have

$$P\{|X - E[X]| \geq k\sigma\} \leq \frac{var(x)}{k^2\sigma^2} = \frac{1}{k^2}$$

Equivalently

$$P\{|X - E[X]| \leq k\sigma\} \geq 1 - 1/k^2$$ where σ is the s.d. of X

Application

Problem

Suppose X is $B:(100, \frac{1}{2})$ random variable. Find a lower bound for the probability that X is between 30 and 70.

$E[X] = np = 100 \times \frac{1}{2} = 50$
$var(x) = npq = 100 \times \frac{1}{2} \times \frac{1}{2} = 25$
$s.d.(x) = \sqrt{25} = 5$

thus, $P(30 \leq X \leq 70) = P\{|X - E[X]| \leq 20\} = P\{|X - E[X]| \leq 4 s.d.(x)\}$
$\geq 1 - \frac{1}{16} \approx 0.94$

Another Example

Suppose Y has an unknown distrubution, with $E[Y] = 50$, and $var(Y) = 25$. Estimate the probability that Y is between 30 and 70.

- Using Chebyshev's enequality

$$P\{30 \leq Y \leq 70\} = P\{|Y - E[Y]| \leq 20\}$$
$$= P\{|Y - E[Y]| \leq 4 s.d.(Y)\} \geq 1 - \frac{1}{16} \approx 0.94$$

- Actual distribution irrelevant. Same bound for all random variables with given mean and variance.