Markov's and Chebychev's Inequalities (8.2)

All r.v's X, Y are either continuous <u>or</u> discrete.

<u>Theorem 1.</u> If X, Y are r.v's on a sample space S, such that $X \le Y$, then $E(X) \le E(Y)$.

<u>Proof.</u> Assume $Y \ge X \ge 0$. We will use

$$(*)\ E(X) = \int_0^\infty P(X > t)\,dt, \quad E(Y) = \int_0^\infty P(Y > t)\,dt.$$

$$\{X > t\} \subset \{Y > t\} \quad \text{for all } t \ge 0$$

$$\therefore P(X > t) \le P(Y > t) \quad " \quad " \quad "$$

$$\therefore \underbrace{\int_0^\infty P(X > t)\,dt}_{\|} \le \underbrace{\int_0^\infty P(Y > t)\,dt}_{\|} \quad " \quad " \quad ".$$

$$\therefore E(X) \le E(Y) \quad \text{by } (*).$$

In general use the extension of $(*)$ to arbitrary X in Ex. 5.2 on p 227

$$(**)\ E(X) = \int_0^\infty P(X > t)\,dt - \int_0^\infty P(X < -t)\,dt.$$

The proof is the same (essentially). $\square$

<u>Theorem 2.</u> (Markov's Inequality). If $Y \ge 0$, then for any constant $m > 0$,

$$P(Y \ge a) \le E(Y^m)\,a^{-m} \quad \text{for all } a > 0$$

<u>Proof.</u> Fix $a > 0$. (i) $1_{\{Y \ge a\}} \le \dfrac{Y^m}{a^m}$. Case 1. $Y \ge a$. RHS $\ge 1 =$ LHS.
Case 2 $Y < a$ RHS $\ge 0 =$ LHS.

So apply Thm 1: $\underbrace{E(1_{\{Y \ge a\}})}_{\|} \le \underbrace{E\left(\frac{Y^m}{a^m}\right)}_{\|}$

$$P(Y \ge a) \le E(Y^m)\,\frac{1}{a^m}. \ \square$$

Theorem 3. (Chebychev's Inequality) Let X be a rv with

$E(X) = \mu$ and $Var(X) = \sigma^2$. Then

(C.I) $\quad P(|X-\mu| \geq a) \leq \dfrac{\sigma^2}{a^2}$ for all $a > 0$.

Proof. Apply Markov's inequality to $Y = |X-\mu|$ with $m = 2$:

$$P(|X-\mu| \geq a) \leq \frac{E(|X-\mu|^2)}{a^2} = \frac{\sigma^2}{a^2}. \quad \square$$

Remarks. ① If σ^2 is small, we get $P(|X-\mu| < a) \geq 1 - \dfrac{\sigma^2}{a^2}$, ie

$|X-\mu|$ is small with probability near 1 by taking a "small" so that $\dfrac{\sigma^2}{a^2}$ is still "small".

② σ is called the standard deviation of X.
If $a = k\sigma$ for $k > 0$, we get
$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

③ (C.I) holds for all rvs with mean μ and variance σ^2.
So for any particular distribution (eg $N(\mu, \sigma^2)$) it is likely to give a poor bound. (see Eg. 2b on p 390)

E.g. N = number of dishes sold at Vijs restaurant on a given weekday.

Past data suggests $E(N) \approx 100$, $Var(N) \approx 20$.

A slow weekday at Vijs occurs when fewer than 85 dishes are sold.

Give an upperbound for P(next weekday is a slow one)

$$P(N \leq 85) \leq P(|N-100| \geq 15)$$

$$\leq \frac{\sigma^2}{15^2} = \frac{20}{225} = \frac{4}{45} \doteq .08$$

[The actual probability is likely to be much less. Chebychev gives a conservative estimate which holds without making further assumptions on the distribution of N. Sometimes conservative estimates are what's needed.]