Inequalities, Convergence and the Weak Law of Large Number

★ The Markov inequality

⇒ Use a bit of information about a distribution to learn something about probabilities of "extreme events".

Markov inequality: If $X \geq 0$ & $a > 0$

then $P(X \geq a) \leq \frac{E[X]}{a}$

$$E[X] = \int_0^\infty x f_X(x)\,dx \geq \int_a^\infty x f_X(x)\,dx$$

$$\geq a \int_a^\infty f_X(x)\,dx = a\,P(X \geq a)$$

$$\Rightarrow P(X \geq a) \leq \frac{E[X]}{a}$$

OR

Let $Y = \begin{cases} 0 & \text{if } X < a \\ a & \text{if } X \geq a \end{cases}$

$\Rightarrow Y \leq X$

$E[Y] \leq E[X]$

$\Rightarrow a\,P(X \geq a) \leq E[X]$

$\Rightarrow P(X \geq a) \leq E[X]/a$

Date
Page

★ The Chebyshev inequality

⇒ Random variable X, with finite mean μ and variance σ^2.

"If the variance is small, then X is unlikely to be too far from mean"

Chebyshev inequality: $P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$

$P(|X-\mu| \geq c)$

$\Rightarrow P((X-\mu)^2 \geq c^2)$

$\Rightarrow P((X-\mu)^2 \geq c^2) \leq \frac{E[(X-\mu)^2]}{c^2}$ {from markovs inequality}

$\Rightarrow P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$

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

★ The weak Law of Large Number (WLLN)

⇒ $X_1, X_2 \cdots$ are iid {indipendent & identically distributed}

⇒ finite mean μ and variance σ^2.

Sample mean: $M_n = \dfrac{X_1 + \cdots X_n}{n}$

$$E[M_n] = \frac{1}{n}\left(E[X_1] + \cdots E[X_n]\right) = \frac{n\mu}{n} = \mu$$

$$Var(M_n) = \frac{1}{n^2} Var(X_1 + \cdots X_n) = \frac{n\sigma^2}{n} = \sigma^2/n$$

$$P(|M_n - \mu| \geq \epsilon) \leq \frac{Var(M_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow{\lim n \to \infty} 0$$

WLLN: For $\epsilon > 0$

$$P(|M_n - \mu| \geq \epsilon) \Rightarrow 0 \text{ as } n \to \infty$$

Date
Page

★ The pollster's problem

⇒ p: fraction of population that will vote "yes" in a referendum.

⇒ i^{th} (randomly selected) person polled:

$$X_i = \begin{cases} 1 & \text{if yes} \\ 0 & \text{if no} \end{cases}$$

$$M_n = \frac{X_1 + \cdots X_n}{n}$$ fraction of yes in our samp.

$$P(|M_{100,000} - p| \geq 0.01) \leq \frac{\sigma^2}{n\epsilon^2} = \frac{p(1-p)}{10^4 \times 10^{-4}}$$

$$\leq \frac{1}{4}$$

$$\frac{1/4}{n\,10^{-4}} \leq \frac{5}{10^2} \iff n \geq \frac{10^6}{20} = 50,000$$

★ Convergence in Probability

Definition: A sequence Y_n converges in probability to a number a if:

for any $\epsilon > 0$, $\lim_{n\to\infty} P(|Y_n - a| \geq \epsilon) = 0$

Date
Page

$\Rightarrow$ Suppose that $X_n \to a$ $Y_n \to b$ in Probability

(#) If g is continuous, then

$$g(X_n) \to g(a)$$

(#) $X_n + Y_n \to a+b$

$\Rightarrow$ But $E[X_n]$ need not converge to a.

★ Comparing $E[g(x)]$ to $g(E[X])$ (Jensen's inequality)

$\Rightarrow$ Let g be convex

$\Rightarrow$ If $0 \leq P \leq 1$, then {definition of convex function}

$$g(Px + (1-P)y) \leq Pg(x) + (1-P)g(y)$$

$$\boxed{g(E[X]) \leq E[g(x)]}$$

★ Hoeffding's inequality for $P(X_1 + \cdots X_n > na)$

$\rightarrow X_i$: iid

$$\boxed{P(X_1 + \cdots + X_n \geq na) \leq e^{-\frac{na^2}{2}}}$$