Chebyshev's Inequality and Law of Large Number

Ang Man Shun

December 6, 2012

Reference Seymour Lipschutz *Introduction to Propability and Statistics*

1 Chebyshev's Inequality

For a random variable $X(\mu, \sigma)$, given any $k > 0$ (no matter how small and how big it is) , the following Propability inequality always holds

$$\text{Prob}\left(\mu - k\sigma \leq X \leq \mu + k\sigma\right) \geq 1 - \frac{1}{k^2}$$

i.e. given the positive number k , the probability that the random variable X within the range $[\mu - k\sigma, \mu + k\sigma]$ is at least $1 - \dfrac{1}{k^2}$

Proof. Consider σ^2 (recall, $\sigma \geq 0$!)

$$\sigma^2 = Var(X) = \sum (x_i - \mu)^2 f(x_i)$$

Now take away all the terms that $|x_i - \mu| \leq k\sigma$, so
- The number of term in the summation is decreased
- Since $\sum (x_i - \mu)^2$ is monotonic increasing, so fewer term, the smaller the value. (As $f(x_i) \in [0, 1]$)
- The remaining term $|x_i - \mu| > k\sigma \quad \Longrightarrow \quad (x_i - \mu)^2 \geq k^2\sigma^2$
- The remaining $f(x_i)$ are those probability that $|x_i - \mu| > k\sigma \qquad \Longrightarrow \sum_{del} f(x_i) = Pr\left(|X - \mu| > k\sigma\right)$

Thus

$$\sigma^2 = \sum (x_i - \mu)^2 f(x_i) \geq \sum_{\substack{\text{Some}\\\text{delete}}} (x_i - \mu)^2 f(x_i) \geq \sum_{\substack{\text{Some}\\\text{delete}}} k^2\sigma^2 f(x_i)$$

i.e.

$$\sigma^2 \geq \sum_{\substack{\text{Some}\\\text{delete}}} k^2\sigma^2 f(x_i) = k^2\sigma^2 \sum_{\substack{\text{Some}\\\text{delete}}} f(x_i) = k^2\sigma^2 \text{Prob}\left(\Big|X - \mu\Big| > k\sigma\right)$$

$$\Longleftrightarrow \quad \sigma^2 \geq k^2\sigma^2 \text{Prob}\left(\Big|X - \mu\Big| > k\sigma\right)$$

Divide both side with σ^2 , as this is a positive term, so the inequality sign does not change

$$\frac{1}{k^2} \geq \text{Prob}\left(\Big|X - \mu\Big| > k\sigma\right)$$

Prob(A) = 1 - Prob(not A)

$$\frac{1}{k^2} \geq 1 - \text{Prob}\left(\Big|X - \mu\Big| < k\sigma\right)$$

Expand the absolute sign : $|x| < y \iff -y < x < y$

$$\frac{1}{k^2} \geq 1 - \text{Prob}\left(-k\sigma < X - \mu < k\sigma\right)$$

$$\frac{1}{k^2} \geq 1 - \text{Prob}\left(\mu - k\sigma < X < \mu + k\sigma\right)$$

i.e.

$$\text{Prob}\left(\mu - k\sigma \leq X \leq \mu + k\sigma\right) \geq 1 - \frac{1}{k^2}$$

□

2 Law of Large Number

For a population $X(\mu, \sigma)$ that, the sampled mean $\bar{x}$ will get closer to the population mean (the real, true mean) as the number of sample increase

i.e.

$$\lim_{n \mapsto \infty} \text{Prob}\left(\mu - \alpha \leq \bar{X}_n \leq \mu + \alpha\right) = 1 \qquad \forall \alpha \quad \text{(No matter how small)}$$

2.1 The Expected value of Sampled Mean

$$\mu_{\bar{X}_n} = E\left[\bar{X}_n\right] = E\left[\frac{X_1 + .. + X_n}{n}\right] = \frac{E(X_1) + ... + E(X_n)}{n} = \frac{\mu + ... + \mu}{n} = \frac{n\mu}{n} = \mu$$

2.2 The Variance of Sampled Mean

$$Var\left(\bar{X}_n\right) = Var\left(\frac{X_1 + ... + X_n}{n}\right) = \frac{1}{n^2} Var\left(X_1 + ... + X_n\right)$$

$$= \frac{Var(X_1) + ... + Var(X_n)}{n^2} = \frac{\sigma^2 + ... + \sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

2.3 The Weak Law of Large Number

By Chebyshev Inequality

$$\text{Prob}\left(\mu - k\sigma \leq X \leq \mu + k\sigma\right) \geq 1 - \frac{1}{k^2}$$

Put $X(\mu, \sigma) = \bar{X}_n \left(\mu_{\bar{X}_n}, \sigma_{\bar{X}_n}\right)$

$$\text{Prob}\left(\mu_{\bar{X}_n} - k\sigma_{\bar{X}_n} \leq \bar{X} \leq \mu_{\bar{X}_n} + k\sigma_{\bar{X}_n}\right) \geq 1 - \frac{1}{k^2}$$

Let $k\sigma_{\bar{X}_n} = \alpha$

$$\text{Prob}\left(\mu_{\bar{X}_n} - \alpha \leq \bar{X} \leq \mu_{\bar{X}_n} + \alpha\right) \geq 1 - \frac{\sigma^2_{\bar{X}_n}}{\alpha^2} = 1 - \frac{\sigma^2}{n\alpha^2}$$

Take limit

$$\lim_{n \mapsto \infty} \text{Prob}\left(\mu_{\bar{X}_n} - \alpha \leq \bar{X} \leq \mu_{\bar{X}_n} + \alpha\right) \geq \lim_{n \mapsto \infty} \left(1 - \frac{\sigma^2}{n\alpha^2}\right) = 1$$

As Prob ≤ 1

$$\lim_{n \mapsto \infty} \text{Prob}\left(\mu_{\bar{X}_n} - \alpha \leq \bar{X} \leq \mu_{\bar{X}_n} + \alpha\right) = 1$$

As $\mu_{\bar{X}_n} = \mu$

$$\lim_{n \mapsto \infty} \text{Prob}\left(\mu - \alpha \leq \bar{X} \leq \mu + \alpha\right) = 1$$

Or

$$\lim_{n \mapsto \infty} \text{Prob}\left(-\alpha \leq \bar{X} - \mu \leq \alpha\right) = 1$$

$$\lim_{n \mapsto \infty} \text{Prob}\left(\left|\bar{X} - \mu\right| \leq \alpha\right) = 1$$

$$\lim_{n \mapsto \infty} \text{Prob}\left(\left|\bar{X} - \mu\right| > \alpha\right) = 0$$

i.e. No matter how small α is , the propability having error in sampled mean $\bar{X}$ and population mean μ is zero.

That is, not possible to have error, then it means $\bar{X} = \mu$

But the condition is $n = \infty$

So the $\bar{X}$ converge weakly to μ when n is large.

$-END-$