Theory of Probability notes

md4178

December 2 2020

Weak Law of Large Numbers

Definition: Let x_1 and x_2 ... be I.I.D random variables, with E[x_i]=$\mu < \infty$, and Var[x_i]=$\sigma^2 < \infty$. For any $\epsilon > 0$, $|x+y|$

P[$|(\frac{1}{n})\sum_{i=1}^{n} x_i| - \mu >= \epsilon] \rightarrow$ 0, as n $\rightarrow$ ∞.

We need to find some inequalities to prove Weak Law of Large Numbers.

Markov Inequality: If we have a random variable x >= 0, then for any a> 0, P[X>=a] <= $\frac{E[X]}{a}$

proof of Markov Inequality:

$$aP[X >= a] = \int_a^\infty af(x)\,dx$$

$$<= \int_a^\infty af(x)\,dx$$

$$<= \int_0^\infty af(x)\,dx \ = E[X]$$

Therefore,

$$aP[X >= a] <= E[x]$$

and

$$P[X >= a] <= \frac{E[x]}{a}$$

Chebyshevi's Inequality: We have a random variable x with E[x]=$\mu < \infty$ and $Var[x] = \sigma^2 < \infty$. Then if we have k>0,

$$P[|x-\mu| >= k] <= \frac{\sigma^2}{k^2}$$

proof of Chebyshevi's Inequality:

From Markov Inequality, we know

$$P[(x-\mu)^2 >= k^2] <= \frac{E[(x-\mu)^2]}{k^2}$$

and this is equivalent to Chebyshevi's Inequality

$$P[|x-\mu| >= k] <= \frac{\sigma^2}{k^2}$$

Proof of Weak Law of Large Numbers(WLLN): We have

$$E[\frac{1}{n}\sum_{i=1}^{n} x_i] = \mu$$

$$Var[\frac{1}{n}\sum_{i=1}^{n} x_i] = \frac{\sigma^2}{n}$$

Using Chebyshevi's Inequality:

$$P[|\frac{1}{n}\sum_{i=1}^{n} x_i - \mu| >= \epsilon] <= \frac{\frac{\sigma^2}{n}}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} P[|\frac{1}{n}\sum_{i=1}^{n} x_i - \mu| >= \epsilon] = 0$$

Central Limit Theorem: Let x_1 and x_2 ... be I.I.D random variables, with E[x_i]=$\mu < \infty$, and Var[x_i]=$\sigma^2 < \infty$. Then,

$$\frac{\frac{1}{n}\sum_{i=1}^{n} x_i - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sum_{i=1}^{n} x_i - n\mu}{\sigma\sqrt{n}}$$

which converges to N(0,1) as $n \to \infty$

Zoom lecture notes:

lemma:Suppose we have x_1, x_2...random variables with CDFs F_{x_i} and MGFs M_{x_i}, for i=1,2,3... We have x with CDF F_x and MGF M_x. If

$$\lim_{n\to\infty} M_{x_i}(t) = M_x(t)$$

for all t, then

$$\lim_{n\to\infty} F_{x_n}(t) = F_x(t)$$

for all t at which F_x is continuous

Proof of Central Limit Theorem:

Assume $\mu = 0, \sigma^2 = 1$

Let $M(t) = E[e^{tx_i}]$, then

$$E[e^{\frac{tx_i}{\sqrt{n}}}] = M(\frac{t}{\sqrt{n}})$$

Note:(if $\mu = 0, \sigma^2 = 1$, then

$$\frac{\frac{1}{n}\sum_{i=1}^{n} x_i - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\frac{1}{\sqrt{n}}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i)$$

furthermore,
$$E[e^{t\sum_{i=1}^{n}\frac{x_i}{\sqrt{n}}}] = (M(\frac{t}{\sqrt{n}}))^n$$

If we set $L(t) = \log(M(t))$

L(0)= $\log(M(0)) = \log(1) = 0$

$L'(t) = \frac{1}{M(t)}M'(t)$

$L'(0) = \frac{M'(0)}{M(0)} = M'(0) = E[x_i] = 0$

$$L''(t) = \frac{M(t)M''(t) - M'(t)M'(t)}{M(t)^2}$$
$$= \frac{M(t)M''(t) - (M'(t))^2}{M(t)^2}$$
$$L(0) = \frac{M(0)M''(0) - (M'(0))^2}{M(0)^2}$$
$$= \frac{E[x^2] - 0}{1} = E[x^2] = 1$$

To prove CLT, we can just show $\lim_{n\to\infty} n\log(M(\frac{t}{\sqrt{n}})) = \frac{t^2}{2}$ and $n\log(M(\frac{t}{\sqrt{n}})$ is just $nL(\frac{t}{\sqrt{n}})$

$$\lim_{n\to\infty} nL(\frac{t}{\sqrt{n}}) = \lim_{n\to\infty} \frac{L(\frac{t}{\sqrt{n}})}{\frac{1}{n}}$$
$$= \lim_{n\to\infty} \frac{-L'(\frac{t}{\sqrt{n}})\frac{1}{2}\frac{t}{n^{\frac{3}{2}}}}{-\frac{1}{n^2}}$$
$$= \lim_{n\to\infty} \frac{tL'(\frac{t}{\sqrt{n}})}{\frac{2}{\sqrt{n}}}$$
$$= \lim_{n\to\infty} \frac{-tL''(\frac{t}{\sqrt{n}})\frac{1}{2}\frac{1}{n^{\frac{3}{2}}}t}{-\frac{1}{n^{\frac{3}{2}}}}$$
$$= \lim_{n\to\infty} \frac{t^2}{2}L''(\frac{t}{\sqrt{n}})$$
$$= \frac{t^2}{2}$$

We can conclude
$$\lim_{n\to\infty} \log(M(\frac{t}{\sqrt{n}}))^n = \frac{t^2}{n}$$
$$\lim_{n\to\infty} M(\frac{t}{\sqrt{n}}))^n = e^{\frac{t^2}{2}}$$

$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i \to\ N(0,1)$ as $n \to\ \infty$

Another method:

$$\lim_{n\to\infty}(M(\frac{t}{\sqrt{n}}))^n = e^(\frac{t^2}{2})$$

Expand $M(\frac{t}{\sqrt{n}})$ in a Taylor Series

$$M(\frac{t}{\sqrt{n}}) = M(0) + \frac{1}{\sqrt{n}}M'(0)t + \frac{1}{n}M"(0)\frac{t^2}{2} + ... =$$

$$= 1 + \frac{1}{n}\frac{t^2}{2} + ...$$

$$(M(\frac{t}{\sqrt{n}}))^n = (1 + \frac{1}{n}\frac{t^2}{2} + ...)^n$$

We also know that

$$\lim_{n\to\infty}(1 + \frac{t}{n})^n = e^t$$

Example questions:

a) If 10 fair dice are rolled, find the approximate probability that the sum obtained is between 30 and 40, inclusive.

Solutions:

Let X_i denote the value of the ith die, i=1,2,3..10. Since$E(X_i) = \frac{7}{2}$, $Var(X_i) = E[X_i^2] - E[X_i]^2 = \frac{35}{12}$ Central Limit Theorem yields

$$P[29.5 <= X <= 40.5] = P[\frac{29.5 - 35}{\sqrt{\frac{350}{12}}} <= \frac{X - 35}{\sqrt{\frac{350}{12}}} <= \frac{40.5 - 35}{\sqrt{\frac{350}{12}}}] = 0.592$$

b) Let X_i =1,2,...10, be independent random variables, each uniformly distributed over (0,1). Calculate an approximation to $P[\sum_{i=1}^{10} X_i > 6]$

Solutions: Since $E[X_i] = \frac{1}{2}$ and $Var(X_i) = \frac{1}{12}$, we have, by the central limit theorem,

$$P[\sum_{i=1}^{10} X_i > 6] = P[\frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{10(\frac{1}{12})}} > \frac{6 - 5}{\sqrt{10(\frac{1}{12})}}]$$