Random Variables Lecture 9

Chebyshev's Inequality

Proposition: (Markov's Inequality) If $X \geq 0$ is a random variable, for any $\alpha > 0$

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

Proof: $\alpha P(X \geq \alpha) = \alpha \sum_{k:\, x_k \geq \alpha} p(x_k) \leq \sum_{k:\, x_k \geq \alpha} x_k p(x_k)$

$\leq \sum_{k} x_k p(x_k) = E[X].$

Proposition: (Chebyshev's Inequality) If X is a random variable with mean μ and variance σ^2, then for any value $k > 0$ $\quad P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$

Proof: $P(|X-\mu| \geq k) = P(|X-\mu|^2 \geq k^2) \leq \frac{E[(X-\mu)^2]}{k^2} =$

$= \frac{Var(X)}{k^2} = \frac{\sigma^2}{k^2}$

Do you see where Markov's Inequality was used to establish our claim?

Ex. Suppose that it is known that a bakery produces X number of cakes per day, where X is a random variable with mean 50.

(a) What can be said about the probability that today over 75 cakes will be produced?

(b) If the variance is known to equal 25, what can be said about the probability that today's production will be between 40 and 60?

Solution:

(a) $P(X \geq 75) \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}$.

(b) $P(40 \leq X \leq 60) = P(-10 < X-50 < 10) =$

$= P(|X-50| < 10) \leq \frac{25}{10^2} = \frac{1}{4}$.

Weak Law of Large Numbers

Thm: Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having mean $E[X_k] = \mu$. Then for any $\varepsilon > 0$,

$$P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) \longmapsto 0 \text{ as } n \longmapsto \infty.$$

Proof: $E\left[\frac{X_1+\ldots+X_n}{n}\right] = \frac{1}{n}\left(E[X_1]+\ldots+E[X_n]\right)$

$= \frac{1}{n} n\mu = \mu.$

Since the X_k are independent, $Cov(X_k, X_j) = 0$ and

$$Var\left(\frac{X_1+\ldots+X_n}{n}\right) = \frac{1}{n^2} Var(X_1+\ldots+X_n) =$$

$$= \frac{1}{n^2}\left(Var(X_1)+\ldots+Var(X_n)\right) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

By Chebyshev's Inequality,

$$P\left(\left|\frac{X_1+\ldots+X_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2 n}$$

Ex. How many times must we toss a fair coin in order to insure that the sample average of the number of tails is within $\frac{1}{10} = 0.1$ of $\frac{1}{2}$ 99% of the time we perform this experiment?

Solution: Let $X_k = \begin{cases} 1 & \text{if } k^{th} \text{ toss is tail} \\ 0 & \text{otherwise.} \end{cases}$

$\mu = E[X_k] = \frac{1}{2}$

$\sigma^2 = E[X_k^2] - (E[X_k])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}$

By the weak Law of large numbers

$$P\left(\left|\frac{X_1+\ldots+X_n}{n} - \frac{1}{2}\right| > \left(\frac{1}{10}\right)\right) \leq \frac{\frac{1}{4}}{\left(\frac{1}{10}\right)^2 n} = \frac{25}{n} \leq \frac{1}{100}$$

whenever $2{,}500 \leq n$

That is, if a fair coin is tossed 2,500 times, you can be 99% certain that the average number of tails is a number between 0.4 and 0.6.

Please note that the weak law of large numbers gives an upper estimate for the number of trials (or necessary size of sample). Think of this upper bound as a confident statement to the effect that 99% of people are less than 14 ft tall.

Strong Law of Large Numbers

Thm: Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having mean $E[X_k] = \mu$

Then $\lim_{n \to \infty} \frac{X_1 + X_2 + \ldots + X_n}{n} = \mu$ with probability 1.

Proof: We will assume $E[X_k^4] = M < \infty$

Assume first that $\mu = 0$. Let $S_n = X_1 + \ldots + X_n$

Then $S_n^4 = (X_1 + \ldots + X_n)^4$ has terms of the form

$X_k X_j X_i X_h \quad k \neq j \neq i \neq h$ and by independence

$E[X_k X_j X_i X_h] = E[X_k] E[X_j] E[X_i] E[X_h] = (E[X_k])^4$

$= 0$.

Similarly $E[X_k^2 X_j X_i] = 0$. In general, the expected value of any term of the form $X_k^{n_k} X_j^{n_j} X_i^{n_i} X_h^{n_h}$

$(n_k + n_j + n_i + n_h = 4)$ where at least one power $= \underline{1}$

is 0.

Thus only terms of the form X_k^4 or $X_k^2 X_j^2$ will have non zero expected value.

Notice that $(X_1 + \dots + X_n)^4$ will have n terms of the form X_k^4 and $\binom{4}{2}\binom{n}{2}$ terms of the form $X_k^2 X_j^2$.

To see that the latter is true notice that there are $\binom{n}{2}$ variable-pair choices where $k < j$. The smaller index variable was contributed by two of the 4 bundles:

$$\underset{1}{(X_1 + \dots + X_n)}\underset{2}{(X_1 + \dots + X_n)}\underset{3}{(X_1 + \dots + X_n)}\underset{4}{(X_1 + \dots + X_n)}$$

The larger index variable was therefore contributed by the remaining two bundles.

Thus $$E\left[\left(\frac{S_n}{n}\right)^4\right] = \frac{1}{n^4}\left(\sum_{k=1}^{n} E[X_k^4] + \sum_{i \neq j} E[X_k^2 X_j^2]\right)$$

$$= \frac{1}{n^4}\left(n E[X_1^4] + \binom{4}{2}\binom{n}{2} E[X_1^2 X_2^2]\right)$$

$$= \frac{1}{n^4}\left(n E[X_1^4] + 3n(n-1)\left(E[X_1^2]\right)^2\right)$$

Since $0 \leq Var(X_1^2) = E[X_1^4] - \left(E[X_1^2]\right)^2$, Hence $(E[X_1^2])^2 \leq E[X_1^4] = M < \infty$

In particular $E[(\frac{S_n}{n})^4] \leq \frac{1}{n^4}(nM + 3n(n-1)M)$

$$\leq \frac{M}{n^3} + \frac{3M}{n^2}$$

So $$\sum_{n=1}^{\infty} E\left[\left(\frac{S_n}{n}\right)^4\right] < \infty \implies E\left[\left(\frac{S_n}{n}\right)^4\right] \longrightarrow 0$$

(The terms of the series must converge to 0)

Thus $$E\left[\sum_{n=1}^{\infty}\left(\frac{S_n}{n}\right)^4\right] = \sum_{n=1}^{\infty} E\left[\left(\frac{S_n}{n}\right)^4\right] < \infty$$

and this means that with probability 1 $\lim_{n\to\infty} \left(\frac{S_n}{n}\right)^4 = 0$

For otherwise there is a nonzero probability p that

$$\sum_{n=1}^{\infty}\left(\frac{S_n}{n}\right)^4 = \infty \quad \text{and then } E\left[\sum_{n=1}^{\infty}\left(\frac{S_n}{n}\right)^4\right] \geq \infty \cdot p = \infty$$

in contradiction to our prior conclusion that this expected value is finite.

Since $\lim_{n\to\infty}\left(\frac{S_n}{n}\right)^4 = 0$ with probability 1

$\lim_{n\to\infty} \frac{S_n}{n} = 0$ with probability 1.

Finally, if $\mu \neq 0$ simply consider $Y_k = X_k - \mu$

By the work above $\lim_{n\to\infty} \frac{Y_1 + \dots + Y_n}{n} = 0$ with probability 1. Hence $\lim_{n\to\infty} \frac{X_1 + \dots + X_n}{n} = \mu$ with probability 1.

How the Weak and Strong Laws of Large Numbers Different?

The weak and the strong law are statements about independent and identically distributed random variables. They are both statements about the limit of sample averages.

The weak law claims that the deviation from the mean of $\frac{X_1+\ldots+X_n}{n}$ becomes less and less likely, but that these deviations may continue to appear infinitely often.

The strong law asserts that with virtual certainty the terms in the sequence $\frac{S_n}{n} = \frac{X_1+\ldots+X_n}{n}$ eventually become indistinguishable from μ.

The difference is most easily understood by looking at the following sequences:

Ex. Consider the sequence $a_n = \frac{1}{n}$ and $b_n = \begin{cases} 1 & \text{if } n = 2^k \\ 0 & \text{otherwise.} \end{cases}$

a_n:

n	a_n
1	1
2	$\frac{1}{2}$
3	$\frac{1}{3}$
$\vdots$	$\vdots$

b_n:

n	b_n
1	0
2	1
3	0
4	1
$\vdots$	$\vdots$

Terms eventually are so small that all look like 0.

There are always terms that don't look like 0, but they become very rare.

The sequence b_n doesn't converge to 0 because the terms in the right column do not all become arbitrarily close to 0. However the distance between successive 1's grows larger and larger.

These ideas are best understood through the language of measure theory.