Limit Theorems
Outline

- Inequalities
- Convergence in Probability
- The Law of Large Numbers
- The Central Limit Theorem
- Refinements

Markov’s Inequality

Theorem. If $Y$ is any RV and $0 < c < \infty$, then
$$P[|Y| \geq c] \leq \frac{1}{c} E[|Y|].$$

Proof. Let $$B = \{|Y| \geq c\}. $$
Then$$ c1_B \leq |Y|.$$So,$$E|Y| \geq E(c1_B) = cP(B).$$

Chebyshev’s Inequality

Corollary. If $X$ has mean and variance
$$\mu = E(X),$$
$$\sigma^2 = E[(X - \mu)^2],$$
then
$$P[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2}.$$Proof. Let
$$Y = |X - \mu|^2$$
in Markov’s Inequality. Then
$$P[|X - \mu| \geq c] = P[Y \geq c^2]$$
$$\leq \frac{1}{c^2} E[Y]$$
$$= \frac{\sigma^2}{c^2}.$$Alternative Statement: Let
$$X^* = \frac{X - \mu}{\sigma}.$$Then
$$E(X^*) = 0,$$
$$D^2(X^*) = 1,$$and
$$P[|X^*| \geq c] \leq \frac{1}{c^2}.$$Remark: General, but seldom sharp.

Example: If $X \sim \text{Normal}[\mu, \sigma^2]$, then $X^* \sim \Phi$, and
$$P[|X^*| \geq 2] = \cdots = .046,$$from the normal tables. Chebyshev asserts (only)
$$P[|X^*| \geq 2] \leq \frac{1}{4}.$$
Bernstein's Inequality

**Corollary.** If $X$ has MGF $M$, then

$$P[X \geq c] \leq \min_{t>0} e^{-ct} M(t).$$

**Proof.** Letting $Y = e^{tX}$ in Markov's Inequality,

$$P[X \geq c] = P[e^{tX} \geq e^{ct}]$$

$$\leq \frac{1}{e^{ct}} E(e^{tX})$$

$$= e^{-ct} M(t)$$

for all $t > 0$. So,

$$P[X \geq c] \leq \min_{t>0} e^{-ct} M(t).$$

---

Example

If $X \sim \Phi$, then

$$M(t) = e^{t^2/2}.$$ 

So,

$$\min_{t>0} e^{-ct} M(t) = \min_{t>0} e^{-ct + t^2/2} = e^{-c^2/2}.$$

For $c = 3$,

$$e^{-3^2/2} = e^{-4.5} = .011 \text{ (Bernstn),}$$

$$1/c^2 = 1/9 = .111 \text{ (Cheb),}$$

$$1 - \Phi(c) = .0013 \text{ (Actual).}$$

---

Sums of Independent RVs

If

$X_1, \ldots, X_n$ are independent,

$$E(X_i) = \mu_i,$$

$$D^2(X_i) = \sigma_i^2$$

and

$$S = X_1 + \cdots + X_n,$$

then

$$E(S) = \mu_1 + \cdots + \mu_n,$$

$$D^2(S) = \sigma_1^2 + \cdots + \sigma_n^2.$$  

**Special Case:** If $\mu_i = \mu$ and $\sigma_i^2 = \sigma^2$ for $i = 1, \ldots, n$, then

$$E(S) = n\mu,$$

$$D^2(S) = n\sigma^2,$$

$$D(S_n) = \sigma \sqrt{n}.$$

---

Special Case: Continued. Let

$$\bar{X} = \frac{S}{n} = \frac{X_1 + \cdots + X_n}{n}.$$  

Then

$$E(\bar{X}) = \frac{1}{n} E(S) = \frac{1}{n} n\mu = \mu,$$

and

$$D^2(\bar{X}) = \left(\frac{1}{n}\right)^2 D^2(S)$$

$$= \frac{1}{n^2} n \sigma^2$$

$$= \frac{\sigma^2}{n}.$$  

Note: $D^2(\bar{X}) \to 0$ as $n \to \infty$. So, for any $\epsilon > 0$,

$$P[|\bar{X} - \mu| \geq \epsilon] \leq \frac{1}{\epsilon^2} D^2(\bar{X})$$

$$= \frac{\sigma^2}{n \epsilon^2}$$

$$\to 0$$

as $n \to \infty$. 

An Example

If

\[ X_1, \ldots, X_n \text{ are independent} \]

and

\[ X_i = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } q = 1 - p \end{cases} \]

then

\[ E(X_i) = p, \]
\[ D^2(X_i) = pq \leq \frac{1}{4}, \]

and, therefore,

\[ E(\bar{X}) = p, \]
\[ D^2(\bar{X}) = \frac{pq}{n} \leq \frac{1}{4n}. \]

Question: Using Chebyshev, find an \( N \) for which

\[ P[|\bar{X} - p| \geq .01] \leq .01 \]

for all \( n \geq N \). Here

\[ P[|\bar{X} - p| \geq .01] \leq \frac{pq}{n(01)^2} \leq \frac{2500}{n} \leq .01 \]

iff

\[ n \geq \frac{2500}{.01} = 250,000. \]

Note: Conservative estimate.

Type of Convergence

If \( Y_n \) are RVs and \( c \) is a constant, then \( Y_n \) converges to \( c \) in mean square iff

\[ \lim_{n \to \infty} E[(Y_n - c)^2] = 0; \]

and \( Y_n \) converges to \( c \) in probability iff

\[ \lim_{n \to \infty} P[|Y_n - c| \geq \epsilon] = 0 \]

for all \( \epsilon > 0 \).

A Simple Relation: If \( Y_n \to^{\text{m.s.}} c \), then \( Y_n \to^{\text{p}} c \), since

\[ P[|Y_n - c| \geq \epsilon] = P[|Y_n - c|^2 \geq \epsilon^2] \leq \frac{1}{\epsilon^2} E|Y_n - c|^2, \]

by Markov’s Inequality.

The Law of Large Numbers

Theorem. Let \( F \) be a distribution function with mean and variance

\[ \mu = \int_{-\infty}^{\infty} x dF(x), \]
\[ \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 dF(x); \]

and let

\[ X_1, \ldots, X_n \sim^{\text{ind}} F. \]

Then

\[ \bar{X}_n = \frac{X_1 + \cdots + X_n}{n} \to \mu \]

in mean square and in probability.

Proof. Here

\[ E[(\bar{X}_n - \mu)^2] = D^2(\bar{X}_n) = \frac{\sigma^2}{n} \to 0, \]

as \( n \to \infty \).
Paraphrase: For repeated trials. Let
\[ \text{Time Ave } = \frac{X_1 + \cdots + X_n}{n} \]
and
\[ \text{Space Ave } = \mu = \int_{-\infty}^{\infty} x f(x). \]
Then
\[ \text{Time Ave } = \text{Space Ave}. \]

Note: Can predict long run behavior.

Example
Roulette

A Single Game: The expected gain is
\[ X = \begin{cases} 
1 & \text{w.p. } 9/19 \\
-1 & \text{w.p. } 10/19 ,
\end{cases} \]
then
\[ E(X) = \frac{9}{19} - \frac{10}{19} = -\frac{1}{19}. \]

Many Games: The actual average gain per play is
\[ \frac{X_1 + \cdots + X_n}{n} \rightarrow -\frac{1}{19}. \]

Note: Qualitative result; quantified later.

---

Indicators

If \( A_1, A_2, \cdots \) are independent with
\[ P(A_i) = p, \]
then
\[ E(1_{A_i}) = p, \]
and
\[ \frac{1}{n}[1_{A_1} + \cdots + 1_{A_n}] \rightarrow p. \]

Note: This is the frequentist interpretation of “probability.”

The Law of Averages: Let
\[ N_n = 1_{A_1} + \cdots + 1_{A_n}. \]
Then
\[ P[A_{n+1} | N_n = k] = p \]
for all \( k \) (assuming that \( p \) is fixed).

The Central Limit Theorem

Recall: If \( F \) is a DF with mean \( \mu \) and variance \( \sigma^2 \),
\[ X_1, \ldots, X_n \sim \text{ind } F, \]
\[ S_n = X_1 + \cdots + X_n, \]
then
\[ E(S_n) = n\mu, \]
\[ D^2(S_n) = n\sigma^2. \]

Let
\[ S^*_n = \frac{S_n - E(S_n)}{D(S_n)} = \frac{S_n - n\mu}{\sigma\sqrt{n}}. \]

Also, let
\[ \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt. \]

The Central Limit Theorem. For all real \( z \),
\[ \lim_{n \to \infty} P[S^*_n \leq z] = \Phi(z). \]
That is
\[ P[S_n \leq x] \approx \Phi(z) \]
for large \( n \). Equivalently,
\[ P[S_n \leq x] \approx \Phi(\frac{x - \eta \mu}{\sigma \sqrt{n}}). \]

**Remarks**

a). (*) is true for *any* \( F \).
b). Speed of convergence does depend on \( F \).
c). Importance of the normal distribution.

---

**Example**

Let
\[ T = \text{Tax}, \]
\[ \langle T \rangle = \text{closest integer}, \]
\[ X = T - \langle T \rangle. \]

Suppose
\[ X \sim \text{Unif}(\frac{1}{2}, \frac{1}{2}) \]

Then
\[ \mu = E(X) = 0, \]
\[ \sigma^2 = \text{Var}(X) = \frac{1}{12}. \]

Let
\[ n = 12,000,000, \]
\[ X_1, \ldots, X_n \sim \text{ind Unif}\left(\frac{1}{2}, \frac{1}{2}\right), \]
\[ S = X_1 + \cdots + X_n. \]

---

**Question:** How big is \( S \)?

**Worst Case Analysis:** \(|S| \leq 6,000,000\), but this is very conservative.

**Probabilistic Analysis:** By CLThm,
\[ S \approx \text{Normal} [n\mu = 0, n\sigma^2 = 1,000,000]. \]

Note:
\[ \sqrt{n\sigma^2} = 1000. \]

So,
\[ P[-2000 \leq S \leq 2000] \]
\[ \approx \Phi\left(\frac{2000 - 0}{1000}\right) - \Phi\left(\frac{-2000 - 0}{1000}\right) \]
\[ = \Phi(2) - \Phi(-2) \]
\[ = .954. \]