Theorem: Markov's inequality

For any $c, r > 0$,

$$P(|X| \geq c) \leq \frac{E[|X|^r]}{c^{(r)}}.$$

Proof

$$P(|X| \geq c) \Leftrightarrow P(|X|^r \geq c^r) \leq \frac{E[|X|^r]}{c^r}.$$

by generalized Markov's inequality. □

Theorem: Chebyshev's inequality

Suppose X is a r.v. with mean $\mu = E[X]$, $\sigma^2 = Var[X]$. For any $k > 0$, then

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

or

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

Proved by Markov:

Proof

$$\begin{aligned} P(|X - \mu| \geq k\sigma) = P((X-\mu)^2 \geq k^2\sigma^2) &\leq \frac{E(X-\mu)^2}{k^2\sigma^2} \\ &= \frac{\mathrm{Var}[X]}{k^2\sigma^2} \\ &= \frac{\sigma^2}{k^2\sigma^2} \\ &= \frac{1}{k^2} \end{aligned}$$

□

Theorem: Jensen's inequality

If g is a convex function, then

$$g(E[X]) \leq E[g(X)].$$

If g is concave, $-g$ is convex. The sign thus flips.

Prove by picture. Compare the tangent and graph of $g(\mu)$.

Theorem: Holder's inequality

Take $p, g \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ (Holder conjugates). Then

$$E[|XY|] \leq (E[|X|^p])^{\frac{1}{p}} \cdot (E[|Y|^q])^{\frac{1}{q}}.$$

We allow $p = 1$ and $q = \infty$ or vice versa.

6. Law of Large Numbers

Definition: distribution of r.v.

The **distribution** of a r.v. X is the probability measure on $\mathbb{R}$, denoted by $P_X(\cdot)$, defined $\forall\ A \subseteq \mathbb{R}$ as

$$P_X(A) = P(\{\omega : X(\omega) \in A\}).$$

Show that this is a probability measure. (empty set is 0, countable additivity, outputs $[0, 1]$).

Theorem: Strong Law of Large Numbers (SLLN)

Suppose that (X_n) is a sequence of independent and identically distributed r.v. (i.i.d.) with $\mu = E[X_n]$ and $E[X_n^4] < \infty$. Then

$$\overline{X_n} := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} \mu.$$

i.e.

$$P\left(\lim_{n\to\infty} \overline{X_n} = \mu\right) = 1.$$

Note. The last condition is always true for simple r.v.

Proof

Transform X so that $\mu = 0$. Consider $E[S_n^4]$. Most terms are just 0, except for

- $E[X_i^2 X_j^2] = (\sigma^2)^2$. There are $3n(n-1)$ choices for 4 different sets of indices.
- $E[X_i^4] < C < \infty$. There are n choices.

So

$$\begin{aligned}
E[S_n^4] &\le C \cdot n + 3n(n-1)(\sigma^2)^2 \\
&= C \cdot n + 3n^2(\sigma^2)^2 - 3n(\sigma^2)^2 \\
&\le C \cdot n + 3n^2(\sigma^2)^2 \\
&\le C \cdot n^2 + 3n^2(\sigma^2)^2 \\
&= kn^2 \qquad \text{where } k = C + 3(\sigma^2)^2 < \infty
\end{aligned}$$

Let $\varepsilon > 0$, use generalized Markov inequality with $g(x) = x^4$,

$$\begin{aligned}
P(|S_n| \ge n\varepsilon) &= P(|S_n|^4 \ge n^4\varepsilon^4) \\
&= P(S_n^4 \ge n^4\varepsilon^4) \\
&\le \frac{E[S_n^4]}{n^4\varepsilon^4} \\
&\le \frac{kn^2}{n^4\varepsilon^4} \\
&= \frac{k}{n^2\varepsilon^4}
\end{aligned}$$

So by Chebyshev,

$$\begin{aligned}
P(|\overline{X_n} - 0| \ge \varepsilon) &= P(|\overline{X_n}| \ge \varepsilon) \\
&= P(|S_n| \ge n\varepsilon) \\
&\ge \frac{k}{n^2\varepsilon^4} \to 0 \text{ as } n \to \infty
\end{aligned}$$

Hence, $\overline{X_n} \xrightarrow{p} 0$.

Note that
$$\sum_{n=1}^{\infty} P(|S_n| \ge n\varepsilon) \ge \frac{k}{\varepsilon^4} \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

By Borel-Cantelli (i), $P(\limsup_n A_n) = 0$ or $P(A_n \ i.o.) = 0$, or $P(|S_n| \ge n\varepsilon \ i.o.) = 0 \Rightarrow P(|\overline{X_n}| \ge \varepsilon \ i.o.) = 0$ for all $\varepsilon > 0$. This is equivalent to

$$P(\lim_{n\to\infty} \overline{X_n} = 0) = 1.$$

Thus, $\overline{X_n} \xrightarrow{a.s.} 0$.

$\square$

Note. The weak law is just convergence in probability instead.

Measure

Definition: Borel sets

In $\mathbb{R}^n$, the σ-field generated by the open rectangles $\{(x_1, \ldots, x_n) : a_i < x_i < b, i = 1, 2, \ldots, n\}$ is called the **Borel sets on $\mathbb{R}^n$**, denoted by $\mathcal{B}(\mathbb{R}^n)$.

Definition

Let $\mathcal{A}$ be a class/collection of sets in Ω. Let $\Omega_0 \subseteq \Omega$ be a set of points.

$$\mathcal{A} \cap \Omega_0 \coloneqq \{A \cap \Omega_0 : A \in \mathcal{A}\}.$$

Note. This is another collection of sets of points.

Theorem: 10.1

Let $\Omega_0 \subseteq \Omega$.

(i) $\mathcal{F}$ is a σ-field on $\Omega \Rightarrow \mathcal{F}_0 \coloneqq \mathcal{F} \cap \Omega_0$ is a σ-field on Ω_0.

(ii) If $\mathcal{F} = \sigma(\mathcal{A})$ on $\Omega \Rightarrow \mathcal{F}_0 \coloneqq \mathcal{F} \cap \Omega_0 = \sigma(\mathcal{A} \cap \Omega_0)$.

Example. $\omega = \mathbb{R}, \Omega_0 = [0, 1]$. This theorem implies that $\mathcal{B}([0, 1]) = \mathcal{B}(\mathbb{R}) \cap [0, 1]$.

Proof

(i) By definition.

(ii) By double containment.

$\subseteq$: define

$$\mathcal{G} = \{A \subseteq \Omega : A \cap \Omega_0 \in \sigma(\mathcal{A} \cap \Omega_0)\}.$$

and want to show $\mathcal{F} \subseteq \mathcal{G}$.

Claim. $\mathcal{A} \subseteq \mathcal{G}$.

$A \in \mathcal{A} \Rightarrow A \cap \Omega_0 \in \mathcal{A} \cap \Omega_0 \subseteq \sigma(\mathcal{A} \cap \Omega_0)$.

Claim. $\mathcal{G}$ is a σ-field on Ω.

The goal is to show that $\mathcal{F}$ is the smallest σ-field containing $\mathcal{A}$, so $\mathcal{G}$ as another σ-field containing $\mathcal{A}$ must contain $\mathcal{F}$.

(i) $\Omega \cap \Omega_0 = \Omega_0 \in \sigma(\mathcal{A} \cap \Omega_0)$ since it is a σ-field on Ω_0.

(ii) Take $A \in \mathcal{G}$,

$$\begin{aligned}(\Omega \setminus A) \cap \Omega_0 &= \Omega_0 \setminus (A \cap \Omega_0) \text{ by a picture} \\ &\in \sigma(\mathcal{A} \cap \Omega_0) \text{ as the complement in } \Omega_0\end{aligned}$$

(iii) Take $A_1, A_2, \ldots \in \mathcal{G}$.

$$\left(\bigcup_{n=1}^{\infty} A_n\right) \cap \Omega_0 = \bigcup_{n=1}^{\infty} (A_n \cap \Omega_0) \in \sigma(\mathcal{A} \cap \Omega_0).$$

$\supseteq$: $\mathcal{A} \cap \Omega_0 \in \sigma(\mathcal{A}) \cap \Omega_0$ since $\mathcal{A} \subseteq \sigma(\mathcal{A})$. Then $\mathcal{A} \cap \Omega_0 \in \mathcal{F} \cap \Omega_0 = \mathcal{F}_0$ which is a σ-field on Ω_0 by part (i). So as the smallest σ-field containing $\mathcal{A} \cap \Omega_0$, $\sigma(\mathcal{A} \cap \Omega_0) \subseteq \mathcal{F}_0 = \sigma(\mathcal{A}) \cap \Omega_0$.

□

Definition

Suppose $(\Omega, \mathcal{F})$. A **general measure** $\mu : \mathcal{F} \to [0, \infty]$ satisfies

(i) $\mu(\emptyset) = 0$.

(ii) countable additivity of disjoint $A_1, A_2, \ldots \in \mathcal{F}$.

Definition: sigma-finite

The measure space $(\Omega, \mathcal{F}, \mu)$ is a **σ-finite space** if Ω can be written as a countable union of $\mathcal{F}$-sets, $A_1, A_2, \ldots$ (not necessarily disjoint), with $\mu(A_n) < \infty$ for all n. Then we say μ is σ-finite.

Note. finite measure $\Rightarrow$ σ-finite.

σ-finite $\not\Rightarrow$ finite.

Example. $(\mathbb{R}, \mathcal{B}, \lambda)$, $\lambda(\mathbb{R}) = \infty$. But $\mathbb{R} = \bigcup_{n=-\infty}^{\infty} [n, n+1) \Rightarrow \lambda([n, n+1)) = 1 < \infty$.

Definition

μ is **concentrated** on $A \in \mathcal{F}$ if $\mu(A^c) = 0$.

Note. μ is concentrated on the support of μ.

Definition: discrete measure

A measure μ is **discrete** if Ω is discrete and if, for any $A \in \mathcal{F}$, $\mu(A) = \sum_{\omega \in A} \mu(\{\omega\})$.

0.1 properties of a general measure μ

1) monotone: $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$. Since $\mu(B) = \mu(A) + \mu(B \setminus A)$.
2) finite subadditivity.
3) countable subadditivity.
4) continuity from below: $A_1, A_2, \ldots \in \mathcal{F}, A \in \mathcal{F}$, then $A_n \uparrow A \Rightarrow \mu(A_n) \uparrow \mu(A)$.
5) continuity from above: $A_n \downarrow A$ and $\mu(A_1) < \infty$, then $\mu(A_n) \downarrow \mu(A)$.

   Example. $\mu(A_1) = \mu(A_2) + \mu(A_1 \setminus A_2) \Rightarrow \mu(A_1 \setminus A_2) = \mu(A_1) - \mu(A_2)$, which only makes sense if $\mu(A_1) < \infty$. Then

   $\mu(A_1) - \mu(A_n) = \mu(A_1 \setminus A_n) \uparrow \mu(A_1 \setminus A)$ by 4. $= \mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$.

Theorem: inclusion-exclusion

$$\mu\left(\bigcup_{i=1}^{n} A_n\right) = \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j}^{n} \mu(A_i \cap A_j) + \ldots + (-1)^{n-1} \mu(A_1 \cap \ldots \cap A_n)$$

Theorem: 10.3

Let $\mathcal{P}$ be a π-system. Suppose that μ_1 and μ_2 are two measures on $\sigma(\mathcal{P})$ that are σ-finite on $\mathcal{P}$ and agree on $\mathcal{P}$, then they agree on $\sigma(\mathcal{P})$.

Proof

Let $A \in \mathcal{P}$. Define

$$\mathscr{L}_A \coloneqq \{B \in \sigma(\mathcal{P}) : \mu_1(A \cap B) = \mu_2(A \cap B)\}.$$

Claim. $\mathcal{P} \subseteq \mathscr{L}_A$.

Take any $B \in \mathcal{P}$. Then $A \cap B \in \mathcal{P}$ by π-system. Then $\mu_1(A \cap B) = \mu_2(A \cap B) \Rightarrow B \in \mathscr{L}_A$.

Claim. $\mathscr{L}_A$ is a λ-system.

Since $\mathcal{P} \subseteq \mathscr{L}_A$, by Dynkin's Theorem, $\sigma(\mathcal{P}) \subseteq \mathscr{L}_A$.

μ_1, μ_2 are σ-finite on $\mathcal{P} \Rightarrow \exists\, A_1, A_2, \ldots \in \mathcal{P}$ such that $\mathcal{P} = \bigcup_{n=1}^{\infty} A_n$ and $\mu_1(A_n) = \mu_2(A_n) < \infty$. Then by inclusion-exclusion,

$$\mu_\alpha \left(\bigcup_{i=1}^{n} (A_i \cap B) \right) = \ldots$$

for $\alpha = 1, 2$. Take any $B \in \sigma(\mathcal{P})$. Then $B \in \mathscr{L}_A$ by $\sigma(\mathcal{P}) \subseteq \mathscr{L}_A$. Since the intersections of A_i is in $\mathcal{P}$ as it is a π-system, this implies that the RHS of inclusion-exclusion agree for $\alpha = 1, 2$. Then LHS also agree:

$$\mu_1 \left(\bigcup_{i=1}^{n} (A_i \cap B) \right) = \mu_2 \left(\bigcup_{i=1}^{n} (A_i \cap B) \right).$$

Denote the union as C_n. Since A_n cover $\mathcal{P}$, and $C_n \uparrow B$, by continuity from below, TODO $\mu_1(B) = \mu_2(B)$. $\square$