Definition: simple function

$f : \Omega \to \mathbb{R}$ is a **simple function** if it can be written as

$$f(\omega) = \sum_{i=1}^{n} x_i I_{A_i}(\omega)$$

where $A_1, \ldots, A_n$ form a partition of Ω (assuming disjoint).

Proposition

$f : \Omega \to \mathbb{R}$ is measurable iff $A_i \in \mathcal{F}$ for $i = 1, 2, \ldots, n$.

Theorem

If $f : \Omega \to \mathbb{R}$ is measurable, there exists a sequence (f_n) of simple measurable functions such that

$$\begin{aligned} 0 \leq f_n(\omega) \nearrow f(\omega) \text{ if } f(\omega) \geq 0 \\ 0 \geq f_n(\omega) \searrow f(\omega) \text{ if } f(\omega) < 0 \end{aligned}$$

Proof

Break up the y-axis at $0, \frac{1}{2^n}, \ldots, \frac{n2^n}{2^n}$ and the negatives of all these. Define

$$f_n(\omega) = \begin{cases} -n & \text{if } f(\omega) \leq -n \\ -(k-1)2^{-n} & \text{if } -k2^{-n} \leq f(\omega) \leq -(k-1)2^{-n} \\ (k-1)2^{-n} & \text{if } (k-1)2^{-n} \leq f(\omega) \leq k2^{-n} \\ n & \text{if } f(\omega) \geq n \end{cases}$$

Clearly f_n are simple functions that are measurable since f is measurable and the interval is a Borel set so the inverse image is in $\mathcal{F}$. $\square$

Intuition. The idea is that $f_n(\omega)$ will expand until its range totally covers $f(\omega)$. Then after that the intervals will just get refined as $n \to \infty$.

Transformations of Measures

Theorem

$(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$, $T : \Omega \to \Omega'$ measurable F/F'. Let μ be a measure on $\mathcal{F}$. Define a function for $\mathcal{F}'$ to $\mathbb{R}$, call it μT^{-1}, as

$$\mu T^{-1}(A') = \mu(T^{-1}(A')) \; \forall \; A' \in \mathcal{F}'.$$

Then μT^{-1} is a measure on $\mathcal{F}'$.

Note. This is the measure-theoretic version of probability distribution transformation using the Jacobian.

14: Distribution Functions

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a r.v.

Definition: distribution

The **distribution** (law) of X is the probability measure P_X, defined as

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

Definition: distribution function

The **distribution function** for X is denoted/defined as

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x).$$

Property (F_X).

1) Non-decreasing: from monotonicity of P or P_X.
2) Right-continuous: from continuity from above for P_X.
3) $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$.
4) F_X can have, at most, countably many points of discontinuity.

Proposition

If a measure μ is σ-finite on a space $(\Omega, \mathcal{F})$, then $\mathcal{F}$ cannot contain an uncountable collection of disjoint sets with positive μ measure.

Intuition. How does this relate to (4)? Let $x = a$ be a point of discontinuity of F_X.

$$P_X(\{x\}) = P_X((-\infty, a]) - P_X((-\infty, a)) = F(a) - F(a^-) > 0.$$

So we see that proving this claim gives us (4), since discontinuities are disjoint singleton sets with positive μ measure.

Proof

Suppose that $\{A_s\}_{s\in S}$ is a disjoint collection of sets each with positive μ measure. We want to show that S must be countable.

Since μ is σ-finite, there exists a cover $\{B_n\}$ of Ω of $\mathcal{F}$-sets with $\mu(B_n) < \infty\ \forall\ n$. For a given n, consider the set of indices:

$$S_n = \{s \in S : \mu(A_s \cap B_n) > 0\}.$$

Claim. For each n, S_n is countable.

Let A be any set in $\mathcal{F}$ ($\mathcal{F}$-set) with $\mu(A) < \infty$. Fix $\varepsilon > 0$, suppose that $s_1, s_2, \ldots, s_k$ are distinct indices such that

$$\mu(A_{s_i} \cap A) \geq \varepsilon > 0.$$

Then

$$k\varepsilon \leq \sum_{i=1}^{k} \mu(A_{s_i} \cap A) \leq \mu(A) \text{ sum of measure of disjoint sets, might not cover } A$$
$$k \leq \frac{\mu(A)}{\varepsilon} < \infty \text{ by assumption}$$

Therefore, the index set must be finite for each $\varepsilon > 0$ (using the same rationals are dense trick). So apply the claim to B_n and union over all the n,

$$\bigcup_{n=1}^{\infty} S_n \text{ is countable.}$$

Since $\mu(A_s) > 0, \{B_n\}$ is a cover of Ω, there exists a n such that $\mu(A_s \cap B_n) > 0$. *i.e.* $s \in S \Rightarrow s \in S_n$. Thus $S \subseteq \bigcup_{n=1}^{\infty} S_n$ must also be countable. $\square$

Theorem: 14.1

Suppose F is non-decreasing, right-continuous, with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$. Then F is the distribution function of some r.v.

Proof

Motivation (pre measure theory): Suppose X has cdf F that is invertible. Then $F(X) \sim \text{Unif}(0,1)$.

Consider the probability space $(\Omega, \mathcal{F}, P)$ with $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}((0,1))$, and P being the Lebesgue measure.

For intuition, let's first suppose that F is invertible. For $\omega \in \Omega$, define $X(\omega) = F^{-1}(\omega)$. Then this X is the r.v. we seek, because

1) X is a r.v. measurable $\mathcal{F}$.

   Take $x \in \mathbb{R}$, since F is invertible and non-decreasing, F^{-1} is strictly increasing. So

$$\begin{aligned} \{\omega : X(\omega) \leq x\} &= \{\omega : F^{-1}(\omega) \leq x\} \\ &= \{\omega : \omega \leq F(x)\} \text{ by increasing} \\ &= (0, F(x)] \in \mathcal{B}((0,1)) = \mathcal{F} \end{aligned}$$

2) $P(X \leq x) = F(x)$.

$$\begin{aligned} P(X \leq x) &= P(\{\omega : X(\omega) \leq x\}) \\ &= P(\{\omega : F^{-1}(\omega) \leq x\}) \\ &= P(\{\omega : \omega \leq F(x)\}) \\ &= \lambda((0, F(x))) \\ &= F(x) - 0 = F(x) \end{aligned}$$

Now let's define a generalized inverse;

$$F^{-1}(\omega) := \inf\{x : \omega \leq F(x)\}.$$

Intuition. This is sort of the "leftmost" x that makes $F(x)$ greater or equal to ω.

Choose $X(\omega) = F^{-1}(\omega)$. Notice that $\inf\{x : \omega \le F(x)\} \le x$ as long as $\omega \le F(x)$ by definition of infimum. So

$$\begin{aligned}\{\omega : X(\omega) \le x\} &= \{\omega : F^{-1}(\omega) \le x\} \\ &= \{\omega : \inf\{x : \omega \le F(x)\} \le x\} \\ &= \{\omega : \omega \le F(x)\}\end{aligned}$$

Then the same proof applies. □

0.1 Weak Convergence (Convergence in Distribution)

Definition: weak convergence

If $F_n, n = 1, 2, \ldots$ and F are distribution functions, then F_n converges weakly to F if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

at all points of continuity x of F. We say that the corresponding r.v.'s converge in distribution (convergence in law). We denote it as $X_n \xrightarrow{d} X$.

Note. It doesn't mean that the r.v.'s are getting close at all, only their distributions.

Example. Given $X_1, X_2, \ldots$ i.i.d $\sim N(\mu, \sigma^2)$, and $X \sim N(\mu, \sigma^2)$ independent of (X_n). Then $X_n \xrightarrow{d} X$ but X_n doesn't converge to X since they are independent.

Theorem

$$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X.$$

Proof

Let F_n be the distribution function of X_n, and F for X. Take $\varepsilon > 0$. Let x be a point of continuity of F.

$$\begin{aligned}F_n(x) = P(X_n \le x) &= P(\{\omega : X_n(\omega) \le x\}) \\ &= P(\{X_n \le x, X \le x+\varepsilon\} \cup \{X_n \le x, X > x+\varepsilon\}) \\ &= P(X_n \le x, X \le x+\varepsilon) + P(X_n \le x, X > x+\varepsilon) \text{ by disjoint} \\ &\le P(X \le x+\varepsilon) + P(|X_n - X| > \varepsilon) \text{ monotonicity, easy to see from drawing}\end{aligned}$$

The other direction gives

$$\begin{aligned}F(X \le x - \varepsilon) &= P(X \le x - \varepsilon, X_n \le x) + P(X \le x - \varepsilon, X_n > x)\\ &\le P(X_n \le x) + P(|X_n - X| > \varepsilon)\end{aligned}$$

Take $n \to \infty$, $X_n \xrightarrow{p} X \Rightarrow \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$. So

$$\begin{aligned}F(x-\varepsilon) - \lim_{n\to\infty} P(|X_n - X| > \varepsilon) \le \lim_{n\to\infty} F_n(x) \text{ if it exists } &\le F(x+\varepsilon) + \lim_{n\to\infty} P(|X_n - X| > \varepsilon)\\ F(x-\varepsilon) \le \lim_{n\to\infty} F_n(x) &\le F(x+\varepsilon)\end{aligned}$$

Finally, take $\varepsilon \to 0$, since x is a point of continuity of F, by Squeeze Theorem the limit of $F_n(x)$ exists and equals to $F(x)$.

Note. More formally, we would squeeze twice for both $\liminf$ and $\limsup$ and show that they are equal.

□

Note. The converse is false.

Example (counterexample)**.** $X_1, X_2, \ldots$ independent with $P(X_i = \pm 1) = \frac{1}{2}$. Suppose X is independent of the X_i with $P(X = \pm 1) = \frac{1}{2}$. Then $X_n \xrightarrow{d} X$ because they have the same distribution functions. However,

$$\begin{aligned}P(|X_n - X| \ge 2) &= P(|X_n - X| = 2)\\ &= P(X_n = 1, X = -1) + P(X_n = -1, X = 1)\\ &= P()P() + P()P() \text{ by independence}\\ &= \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} \ne 0\end{aligned}$$

Example. X_n with cdfs F_n, where

$$\lim_{n\to\infty} F_n(x) = \begin{cases}0, & x < c\\ 1, & x \ge c\end{cases}$$

So $X_n \xrightarrow{d} X$ where $X = c$ w.p. 1.

Example.

$$\lim_{n\to\infty} F_n(x) \begin{cases}0, & x \le c\\ 1, & x \ge c\end{cases}$$

This is not a valid cdf but matches at all points of continuities. Then $X_n \xrightarrow{d} X$ where $X = c$ w.p. 1.