Limit Thoerems and Finitiely Additive Probability

Rajeeva L. Karandikar

Director

Chennai Mathematical Institute

rlk@cmi.ac.in

rkarandikar@gmail.com

Most of us who study Probability theory in modern times have a belief that the countability additivity axiom introduced by Kolmogorov is a must for limit theorems to hold. Infact, some believe that countability additivity is required for two reasons- for the integration theory and for limit theorems.

We forget that some of the theorems have a much longer history- Law of large numbers and Central limit theorem.

We also forget, or rather are never told, what Kolmogorov himself wrote when he introduced the axiom of countable additivity. In his 1933 monograph, he goes onto give motivation for finite additivity assumption on probability measures.

He then introduces the countably additivity axiom as a regularity assumtion and goes onto say that he will not try to give motivation for it as in reality we can only observe finitely many events at a time.

So we must use as regular a model as is consistent with other requirements but be willing to consider a model without this regularity assumption if some natural requirement is in conflict with the countability additivity assumption.

This is the case when engineers consider white noise which we say does not exist- we mean that white noise does not exist on a countably additive probability space, for if it does, its indefinite inegral would give a Brownian motion with differentiable paths.

Indeed, White Noise can be constructed on a finitely additive probability space and one can do a lot of meaningful analysis with it. Kallianpur and myself have worked on non-linear filtering theory in this framework.

I will leave that for another time and return to Limit Theorem

Let us consider i.i.d. Bernouli (p) random variables $X_1, X_2, \ldots, X_n$ (*i.e.* $\mathbb{P}(X_j = 0) = 1 - p$ and $\mathbb{P}(X_j = 1) = p$ for $1 \leq j \leq n$ and $X_1, X_2, \ldots, X_n$ are independent. The (Weak) Law of large numbers says that for all $\varepsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}(\mid \frac{X_1 + X_2 + \ldots + X_n}{n} - p \mid > \varepsilon) = 0.$$

The formulation as well as proof of this result does not depend upon countably additivity.

The DeMovire Central Limit theorem: Indeed one has for all x

$$\lim_{n\to\infty} \mathbb{P}(\frac{X_1 + X_2 + \ldots + X_n - np}{\sqrt{npq}} \leq x)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-\frac{1}{2}u^2\} du.$$

As far the statement is concerned, one does not need countable additivity and the early proofs of the result use sterlings formula and are valid without bringing in Kolmogorov framework.

Strong Law of Large Numbers:

Let $X_1, X_2, \ldots X_n, \ldots$ be independent random variables with $\mathbb{P}(X_j = 0) = 1 - p$ and $\mathbb{P}(X_j = 1) = p$ for $j \geq 1$ where $0 < p < 1$. Then

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \rightarrow p \;\; a.s$$

equivalently for all $\varepsilon > 0$

$$\lim_{m \to \infty} \mathbb{P}(\sup_{n \geq m} \mid \frac{X_1 + X_2 + \ldots + X_n}{n} - p \mid > \varepsilon) = 0.$$

In this case, even for formulation, one needs to have the infinte sequence of random variables and hence this depends heavily on the Kolmogorov framework.

If $X_1, X_2, \ldots, X_n$ are i.i.d. E valued random variables where E is a finite subset of $\mathbb{R}$. Then one can show that the weak law of large numbers and central limit theorem are true for this sequence. Once again the formulation as well as the proof can be worked out without Kolmogorov's frame work.

Strong law of large numbers seems to need counably additivity to even formulate.

We are going to see that various limit laws are valid on finitely additive measure spaces as well.

Indeed, I will show that *Almost all limit laws that are valid on countably additive probability spaces are also valid on finitely additive probability spaces*

For this we need to go over notion of measurable functions and integration theory w.r.t finitely additive measures with some details.

As I said earlier, one tends to also believe that the integration theory and that the convergence theorems depend upon counable additivity of the underlying measure.

The classic book by - Dunford and Schwartz (Linear Operators -I) devolops integration with respect to a finitely additive measure first, goes onto prove dominated convergence theorem with convergence in measure and then deduce the usual DCT when the measure is countably additive.

Let H be a non empty set, $\mathscr{C}$ be a field of subsets of H and μ be a finitely addtive measure on $(H,\mathscr{C})$. For any set $A \subseteq H$, let

$$\mu^*(A) = \inf\{\mu(C) : A \subseteq C,\ C \in \mathscr{C}\}.$$

As in the case of countably addtitive measure, one can assume that $(H,\mathscr{C},\mu)$ is complete:

$A \subseteq H$, $\mu^*(A) = 0$ implies $A \in \mathscr{C}$.

A function $f : H \mapsto \mathbb{R}$ is said to be simple if

$$f(x) = \sum_{k=1}^{m} a_k \mathbf{1}_{B_k}(x)$$

for some $B_1, \ldots B_m \in \mathscr{C}$ and $a_1, \ldots, a_m \in \mathbb{R}$ and for such a simple function f we define

$$\int f \, d\mu = \sum_{k=1}^{m} a_k \mu(A_k).$$

The class of *measurable functions* $\mathsf{L}^0(H,\mathscr{C},\mu)$ is the class of functions f for which there exist simple fucntions $\{f_n\}$ such that for every $\varepsilon > 0$

$$\mu^*(|f_n - f| > \varepsilon) \to 0.$$

For a measurable function f it can be seen that

$$\{f \leq t\} \in \mathscr{C} \; \forall t \in U_f$$

where U_f^c is at most countable. Indeed, U_f is the set of continuity points of the increasing function $t \mapsto \mu^*(f \leq t)$.

The class of *integrable functions* $\mathsf{L}^1(H, \mathscr{C}, \mu)$ is the class of functions f for which there exist simple fucntions $\{f_n\}$ such that for every $\varepsilon > 0$

$$\mu^*(|f_n - f| > \varepsilon) \to 0$$

and

$$\lim_{m,n\to\infty} \int |f_n - f_m| \, d\mu = 0$$

and for such an f we define

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$$

In this framework we have a version of DCT:

Let $g_n, g \in \mathsf{L}^1(H, \mathscr{C}, \mu)$ be such that for every $\varepsilon > 0$

$$\mu^*(| g_n - g |> \varepsilon) \to 0$$

and that there exists $h \in \mathsf{L}^1(H, \mathscr{C}, \mu)$ such that $\int h \, d\mu < \infty$ and

$$| g_n | \leq h \ \forall n.$$

Then

$$\int | g_n - g | \ d\mu \to 0$$

and

$$\int g_n \, d\mu \to \int g \, d\mu .x$$

Example: A finitely additive measure that is not countably additive:

One can show that on the set of natural numbers N, there esists a measure μ on the power set $\mathscr{P}(\mathsf{N})$ which gives probability 0 to every singleton and 1 to the whole set. Now $f(n) = \frac{1}{n}$ is a measurable and integrable function on $(\mathsf{N}, \mathscr{P}(\mathsf{N}), \mu)$. Note that here $f > 0$ but $\int f \, d\mu = 0$.

Example: Canonical Gauss Measure on a Hilbert Space.

Let H be a real separable Hilbert space and let $\{e_n : n \geq 1\}$ be an orthonormal basis. Let $X_k : H \mapsto \mathbb{R}$ be defined by $X_k(h) = \langle h, e_k \rangle$. Let $\mathscr{C}_n$ denote sets of the form

$$\pi_n^{-1}(B) = \{h : \ (X_1(h), \ldots X_n(h)) \in B\}, \ B \in \mathscr{B}(\mathbb{R}^n).$$

Define μ_n on $\mathscr{C}_n$ by

$$\mu_n(\pi_n^{-1}(B)) = \mathbb{P}((Z_1, Z_2, \ldots, Z_n) \in B)$$

where $\{Z_k\}$ is a sequence of i.i.d. $N(0,1)$ random variables on a countably additive probability space.

It can be shown that $\mathscr{C}_n \subseteq \mathscr{C}_{n+1}$ and that $\{\mu_n\}$ is a consistant family and thus we can define μ on $\mathscr{C} = \cup_{n=1}^{\infty} \mathscr{C}_n$ satisfying

$$\mu(A) = \mu_n(A) \text{ for } A \in \mathscr{C}_n,\ n \geq 1.$$

It may be noted that each $\mathscr{C}_n$ is a σ-field and μ_n is a countably additive measure while $\mathscr{C}$ is a field and μ is not countably additive on the field.

To see this, note that each $n \geq 1$, $X_1, X_2, \ldots, X_n$ are iid $N(0,1)$ and hence

$$\mu(\{\sum_{k=n}^{2n} X_k^2 \geq 1\}) \to 1$$

while

$$\sum_{k=1}^{\infty} X_k(h)^2 = \|h\|^2 < \infty.$$

These two relations would contradict each other if μ is countably additive on $\mathscr{C}$.

If $\mathscr{C}^*$ is the completion of $\mathscr{C}$ under μ, then μ on $(H,\mathscr{C})$ is called the canonical Gauss measure on H.

This measure plays an important role in the theory of White Noise and also in the theory of Abstrat Wiener spaces.

Dubins and Savage framework :

Dubins and Savage introduced a framework for finitely additive probability that they called the Strategic framework.
This framework was the starting point for exploring limit theorems on fintely additive probabiltiy spaces.

(From obituiry for Lester Dubins)

… This encounter with Savage developed into a collaboration generating several key papers and culminating in the ground-breaking monograph *How to Gamble if You Must (Inequalities for Stochastic Processes)*, which presented a coherent theory of gambling processes and optimal behavior in gambling situations. Influenced by Bruno de Finetti, the two collaborators worked in a finitely additive framework in order to bypass the measurability difficulties inherent in maximizing groups constituted of so many functions that they could not be counted.

Dubins and Savage, in order to avoid measurability questions, decided to work with finitely additive measures defined on the power set of the underlying set.

Fact: Any finitely additive measure on $(H,\mathscr{C})$ can be extended as a fintely additive measure to $(H,\mathscr{P}(H))$.

Dubins and Savage proposed a setup where given a distribution σ_1 of X_1 and conditional distribution $\sigma_n(\cdot;a_1,a_2,\ldots,a_{n-1})$ of X_n given $X_1=a_1,\ldots,X_{n-1}=a_{n-1}$, [where $\sigma_1(\cdot)$ and $\sigma_n(\cdot,a_1,a_2,\ldots,a_{n-1})$ are finitely additive measures on $(S,\mathscr{P}(S))$ for $a_1,\ldots,a_k,\ldots\in S$], they showed a cannonical way of constructing a measure on the power set of S^∞ so that the coordinate mappings have required properties.

There is a natural way of defining Independence in this setting.

Purves - Sudderth and Chen proved some limit theorems in the strategic setting of Dubins and Savage for a sequence of Independent random variables and also for martingales. Ramakrishnan proved the Central limit theorem in this setting - he showed that Lindeberg-Feller Central limit theorem is true verbatim in the Dubins Savage framework.

In particular, Ramakrishnan showed that if $X_1, X_2, \ldots$ are iid with mean 0 and variance 1, then

$$\lim_{n\to\infty} \sigma\left(\frac{X_1 + X_2 + \ldots + X_n}{\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-\frac{1}{2}u^2\} du.$$

One of the key techniques used was to approximate the given iid sequence $\{X_n\}$ by a sequence $\{Y_n\}$ such that each Y_n takes finitely many values and Y_n's are independent. Further, Y_n is a function of X_n.

For any m, the distribution of $(Y_1, Y_2, \ldots, Y_m)$ is trivially countably additive. In the Dubins Savage framework, the distribution of the full sequence $\{Y_n : n \geq 1\}$ is countably additive.

Then one uses a limit theorem on $\{Y_n\}$ to get the required result. Of course, as n becomes large, Y_n comes closer to X_n and hence Y_n is no longer iid.

Thus to prove CLT for an iid sequence $\{X_n\}$, one needs to use Lindeberg-Feller CLT for Y_n. Of course, this means one has to manage the approximation suitably. To prove SLLN for an iid sequence $\{X_n\}$, a suitable version of SLLN for an independent sequence Y_n is needed, which is readily available.

However, when trying to prove Law of iterated logarithm for an iid sequence, a version for independent sequences (with different distributions) is needed, which was not readily available.

While Ramakrishnan and I were trying to prove the law of iterated logerithm for an iid sequence and extend CLT to path space to get an analogue of Donsker's invariance principle, (we were nearly there) he mentioned to me that some experts believed that there should be a meta theorem that states that most limits theorems that are true under countability additivity axiom are also ture without it.

How does one formulate such a meta theorem?

When I had been thinking of this, we had a seminar by Prof S D Chatterjee at ISI Kolkata on his Subsequence principle and its proof by Aldous that showed me a way. This subsequence principle is interesting in its own right.

Chatterjee's Subsequence Principle:

Komlos had proven in 1967 that *Let* $X_1, X_2, \ldots, X_n \ldots$ *be any sequence of random variables such that*

$$\sup_{n \geq 1} \mathbb{E}[|X_n|] < \infty.$$

Then there exists a subsequence $\{n_k\}$ *and a random variable* Z *such that*

$$\lim_{k \to \infty} \frac{X_{n_1} + X_{n_2} + \ldots + X_{n_k}}{k} = Z, \ \ a.s.$$

Chaterjee proved a similar version for the law of iterated logarithm:

Let $X_1, X_2, \ldots, X_n \ldots$ be any sequence of random variables such that

$$\sup_{n\geq 1} \mathbb{E}[|X_n|^2] < \infty.$$

Then there exists a subsequence $\{n_k\}$ and random variables U, V such that

$$\limsup_{n\to\infty} \frac{X_{n_1} + X_{n_2} + \ldots + X_{n_k} - kU}{\sqrt{2k \log\log(k)}} = V \ a.s.$$

Chatterjee also proved a version of the subsequence result for CLT:

Let $X_1, X_2, \ldots, X_n \ldots$ be any sequence of random variables such that

$$\sup_{n \geq 1} \mathbb{E}[|X_n|^2] < \infty.$$

Then there exists a subsequence $\{n_k\}$ and random variables U, V such that

$$\frac{X_{n_1} + X_{n_2} + \ldots + X_{n_k} - kU}{\sqrt{k}} \to VZ$$

where Z is a random variable independent of V having standard normal distribution and the convergence above is in distribution.

Based on these results, Chatterjee formulated the following heuristic principle:

Given a limit theorem for i.i.d. random variables under certain moment conditions, there exists an analogous theorem such that an arbitrarily-dependent sequence (under the same moment conditions) always contains a subsequence satisfying this analogous theorem.

The phrase *same moment condition* above has to be interpreted as follows:

The condition $\mathbb{E}[|X_1|^p < \infty$ for the iid sequence $\{X_n\}$ is to be read as

$$[\sup_n \mathbb{E}[|X_n|^p] < \infty$$

and hence the same condition on an arbitrary sequence $\{Y_n\}$ turns out to be

$$[\sup_n \mathbb{E}[|Y_n|^p] < \infty.$$

Aldous formulated a precise result which captures the Chatterjee's subsequence princile and proved the same. This had appeared in 1977. Chatterjee himself presented this in a colloquim talk at ISI in early 1980.

This gave me a way of formulating the elusive meta theorem I had been thinking about.

Back to Finitiely additive measures:

Dubins Savage framework:

Let ν_n be a sequence of finitely additive probability measure on $(\mathbb{R}, \mathscr{P}(\mathbb{R}))$. Taking $\sigma_1 = \nu$ and $\sigma_n(\cdot, a_1, a_2, \ldots, a_{n-1}) = \nu_n$ we construct in the canonical strategic way a finitely additive probability measure σ n $(\mathbb{R}^\infty, \mathscr{P}(\mathbb{R}^\infty))$ such that, denoting the coordinate mappings on $\mathbb{R}^\infty$ by $\{X_n\}$,

$$\sigma(X_n \in A_n, n \geq 1) = \prod_{n=1}^{\infty} \nu_n(A_n)$$

for all $A_n \subseteq \mathbb{R}$, $n \geq 1$.

Thus $\{X_n\}$ are independent r.v.'s with marginal distributions ν_n.

Let us assume that for each n, ν_n is tight- *i.e.* given $\varepsilon > 0$, $\exists K_n < \infty$ such that

$$\nu_n([-K_n, K_n]) \geq 1 - \varepsilon.$$

Note that any moment condition: $\int |X_n|^p \, d\sigma < \infty$ for $p > 0$ would imply tightness for ν_n.

Let $F_n(t) = \sigma(X_n \leq t)$ denote the distribution of X_n.

Now F_n is an increasing function and $\lim_{t\to-\infty} F_n(t) = 0$ and $\lim_{t\to\infty} F_n(t) = 1$ in view of tightness of ν. Hence

$$G_n(t) = F_n(t+) = \lim_{s\downarrow t} F_n(s)$$

is a proper distribution function. Thus we can construct a countably additive probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a sequence of independent r.v.'s $\{Z_n\}$ on it with marginal distribution functions G_n.

Note that

$$F_n(t) = G_n(t) \text{ for all continuity points } t \text{ of } F_n.$$

This leads to $\int h(X_n)\,d\sigma = \int h(Z_n)\,d\mathbb{P}$ for all bounded continuous functions h and for $m \geq 1$ and for all $h \in C_b(\mathbb{R}^m)$

$$\int h(X_1,\ldots X_m)\,d\sigma = \int h(Z_1,\ldots,Z_m)\,d\mathbb{P} \qquad (1)$$

Indeed, (1) is also true if h is positive or if $\int |h(Z_1,\ldots,Z_m)|\,d\mathbb{P} < \infty$.

Thus a moment condition holds for $\{X_n\}$ if and only if it holds for $\{Z_n\}$.

We are going to prove that *a limit theorem holds for* $(X_1,\ldots,X_n,\ldots)$ *if and only if it holds for* $(Z_1,\ldots,Z_n,\ldots)$.

Theeorem: Suppose $A \subseteq \mathbb{R}^\infty$ satisfies for some p, $0 < p < \infty$

$$(x_1, x_2, \ldots) \in A \text{ and } \sum_{n=1}^{\infty} |x_n - y_n|^p < \infty \text{ implies } (y_1, y_2, \ldots) \in A \tag{2}$$

(Such A has been called limit statute by Aldous). Then

$$(X_1, X_2, \ldots) \in A \; a.s. \text{ if and only if } (Z_1, Z_2, \ldots) \in A \; a.s.$$

Example: (SLLN) Fix $\alpha \in \mathbb{R}$. Let

$$A = \{(x_1, x_2, \ldots) \in \mathbb{R}^\infty : \frac{x_1 + \ldots + x_n}{n} \to \alpha\}.$$

This satisfies (2) with $p = 1$. Thus we have

$$\frac{X_1 + \ldots + X_n}{n} \to \alpha \; a.s. \text{ if and only if } \frac{Z_1 + \ldots + Z_n}{n} \to \alpha \; a.s.$$

So if $\{X_n\}$ are iid with $\int X_1 d\sigma = \alpha$ then $\{Z_n\}$ are iid with $\mathbb{E}(Z_1) = \alpha$. So by SLLN, $\frac{Z_1+\ldots+Z_n}{n} \to \alpha \; a.s.$ and hence

$$\frac{X_1 + \ldots + X_n}{n} \to \alpha \; a.s.$$

Let $\{X_n\}$ be iid with $\int X_1 d\sigma = \alpha$ and $\int (X_1 - \alpha)^2 d\sigma = \beta$. Let

$$A = \{(x_1, x_2, \ldots) \in \mathbb{R}^\infty : \limsup_{n\to\infty} \frac{x_1 + \ldots + x_n - n\alpha}{\sqrt{\beta 2n \log\log(n)}} = 1\}$$

Once again easy to see that A satisfies (2) and thus the law of oterated logarithm for the sequence $\{Z_n\}$ (which is now iid with $\mathbb{E}(Z_1) = \alpha$ and $\mathbb{E}(Z_1 - \alpha)^2 = \beta$)) implies that the same is true for $\{X_n\}$.

Indeed, we can likewise show that Strassen's law of iterated logarithm is also true for $\{X_n\}$.

Proof: The idea is to approximate $\{X_n\}$ by $\{Y_n\}$ and $\{Z_n\}$ by $\{W_n\}$ such that each Y_n and W_n take finitely many values and have same distribution. The approximation is such that

$$\sum_{n=1}^{\infty} |X_n - Y_n|^p < \infty \, a.s.$$

$$\sum_{n=1}^{\infty} |Z_n - W_n|^p < \infty \, a.s.$$

and hence $\{X_n\} \in A$ *a.s.* if and only if $\{Y_n\} \in A$ *a.s.* and $\{Z_n\} \in A$ *a.s.* if and only if $\{W_n\} \in A$ *a.s.* (using (2) on A). Then using $\{Y_n\}$ and $\{Z_n\}$ have same distribution we conclude $\{X_n\} \in A$ *a.s.* and $\{Z_n\} \in A$ *a.s.*

Recall F_n, G_n are distribution functions of X_n, Z_n and are equal at all continuity points of F_n. Get $a_{n,0} < a_{n,1} < \dots a_{n,m_n}$ continuity points of F_n such that

$$F_n(a_{n,0}) < \frac{1}{2^n},\ F_n(a_{n,n_m}) > 1 - \frac{1}{2^n}$$

$$a_{n,j+1} - a_{n,j} < \frac{1}{2^n},\ \ 0 \leq j < n_m$$

and let $\phi_n(x) = a_{n,j}$ for $a_{n,j} \leq x < a_{n,j+1}$ and $\phi_n(x) = a_{n,0}$ for $x < a_{n,0}$ and $\phi_n(x) = a_{n,n_m}$ for $x \geq a_{n,n_m}$.

Let $Y_n = \phi_n(X_n)$ and $W_n = \phi_n(Z_n)$. Then for each n Y_n and W_n have same law.

Since $\mathbb{P}(|Z_n - W_n| > \frac{1}{2^n}) \leq \frac{1}{2^n}$, using Borel Cantelli Lemma it follows that

$$\sum_{n=1}^{\infty} |Z_n - W_n|^p < \infty \, a.s.$$

Since the event $|X_n - Y_n| > \frac{1}{2^n}$ depends only on X_n, the Borel Cantelli Lemma holds in the Dubins Savage framework also and so using $\mathbb{P}(|X_n - Y_n| > \frac{1}{2^n}) \leq \frac{1}{2^n}$ it follows that

$$\sum_{n=1}^{\infty} |X_n - Y_n|^p < \infty \, a.s.$$

We have noted that for each n, Y_n and W_n have same law and using independence, we have that for each n
$(Y_1, \ldots, Y_n)$ and $(W_1, \ldots, W_n)$ have same laws.
As remarked earlier, in the Dubins Savage framework, using the fact that Y_n takes finitely many values and that $Y_n = \phi_n(X_n)$, it follows that the law of the sequence $\{Y_n\}$ is countably additive and hence $(Y_1, \ldots, Y_n, \ldots)$ and $(W_1, \ldots, W_n, \ldots)$ have same laws.

Thus

$$\begin{aligned}\{X_n\} \in A \; a.s. &\text{ if and only if } \{Y_n\} \in A \; a.s.\\ &\text{ if and only if } \{W_n\} \in A \; a.s.\\ &\text{ if and only if } \{Z_n\} \in A \; a.s.\end{aligned}$$

Caution: The Almost Sure Limit theorems may not be true on a general finitely probability space.

Convergence in distribution:

Let $\{X_k\}$ be the sequence of independent random variables on $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty), \sigma)$. Let S be a complete separable metric space and let $g_n : \mathbb{R}^\infty \mapsto S$ be given mappings. Let λ be a countably additive probability measure on S. Let $\xi_n = g_n(X_1, X_2, \ldots)$. We say that ξ_n converges in distribution to λ ($\xi_n \xrightarrow{d} \lambda$) if

$$\int h(\xi_n)\, d\sigma \to \int h\, d\lambda$$

for all bounded continuous functions h on S.

Note we have required λ to be countably additive. Otherwise limit may not be unique.

It can be shown here that

$$\int h(\xi_n)\, d\sigma \to \int h\, d\lambda$$

for all bounded continuous functions h on S if and only if for all Borel sets B in S with $\lambda(\partial B) = 0$,

$$\sigma(\xi_n \in B) \to \lambda(B).$$

Let $g_n : \mathbb{R}^\infty \mapsto S$ be a sequence of functions such that

$$|g_n(x_1, x_2, \ldots) - g_n(y_1, y_2, \ldots)| \sum_{j=1}^{\infty} C_{n,j} |x_j - y_j|^p$$

for $0 < p < \infty$, and $c_{n,j}$ are constants such that $0 \leq C_{n,j} \leq 1$ and $\lim_n C_{n,j} = 0$ for each j.

Theorem. Let $\{X_n\}$ and $\{Z_n\}$ be as in previous theorem. Let λ be a counably additive probability measure on S and let $\alpha \in S$. Then

$$g_n(X_1,X_2,\ldots) \xrightarrow{d} \lambda \text{ if and only if } g_n(Z_1,Z_2,\ldots) \xrightarrow{d} \lambda$$

and

$$g_n(X_1,X_2,\ldots) \xrightarrow{\sigma} \alpha \text{ if and only if } g_n(Z_1,Z_2,\ldots) \xrightarrow{\mathbb{P}} \alpha$$

where $\xrightarrow{\sigma}$ means convergence in σ probability and $\xrightarrow{\mathbb{P}}$ means convergence in $\mathbb{P}$ probability.

We approximate $\{X_n\}$ by $\{Y_n\}$ and $\{Z_n\}$ by $\{W_n\}$ as in the earlier case. Recall the distributions of $\{Y_n\}$ and $\{W_n\}$ are the same and it can be shown that

$$g_n((X_1, X_2, \ldots)) - g_n((Y_1, Y_2, \ldots)) \xrightarrow{\sigma} 0$$

and

$$g_n((Z_1, Z_2, \ldots)) - g_n((W_1, W_2, \ldots)) \xrightarrow{\mathbb{P}} 0.$$

Now the proof follows from these observations.

Taking

$$g_n(x_1, x_2, \ldots) = \frac{x_1 + x_2 + \ldots + x_n - n\alpha}{\sqrt{n\beta}}$$

we can deduce the CLT proven by Ramakrishnan: Let $X_1, X_2, \ldots$ be iid r.v.'s in the Dubins Savage framework with $\int X_1 d\sigma = \alpha$ and $\int (X_1 - \alpha)^2 d\sigma = \beta < \infty$.

Let

$$\xi_n = \frac{x_1 + x_2 + \ldots + x_n - n\alpha}{\sqrt{n\beta}}$$

and let λ denote that standard normal distribution on $\mathbb{R}$. Then

$$\xi_n \xrightarrow{d} \lambda.$$

We can deduce the Lindeberg-Feller Central Limit theorem as well with some work as the conditions are in terms of *generalised* moments and we need to show that these can be handled as well- namely Lin.deberg-Feller conditions hold for $\{X_n\}$ if and only if Lindeberg-Feller conditions hold for $\{Z_n\}$. Indeed, we can prove the Donsker's invariance principle as well.

One problem with Dubins Savage framework:

An iid sequence need not be exchangebale

This is because in extension to power sets at each stage, we are making some arbitrary choices and so natural symmetries are not being preserved.

It was OK for Dubins-Savage where the index and conotation of time and hence not being exchangebale did not bother them, but it is not desirable in general.

So I turn my attention to a general finitely additive probability space. And in the same breath let us get over independence as well.

So let $H,\mathscr{C},\mu)$ be any complete finitely additive probabiltiy space and let $X_1,X_2,\ldots X_n,\ldots$ be any $\mathbb{R}$ valued sequence of random variables (not assumed to be independent).

Fix n and consider the mapping L_n for $h \in C_n(\mathbb{R}^n)$

$$L_n(h) = \int h(X_1, \ldots, X_n)\, d\mu.$$

L_n is a positive linear function and it can also be shown that if $h_k \downarrow 0$, $L_n(h_k) \to 0$ as $k \to \infty$. Thus by Daniell's Theorem, there exists a countably additive measure λ_n such that

$$L_n(h) = \int h\, d\lambda_n.$$

Easy to see that λ_n's are consitent and thus determine a countably additive probability measure $\mathbb{P}$ on $\Omega = \mathbb{R}^\infty$ such that writing Z_n to be coordinate mappings on $\mathbb{R}^\infty$

$$\int h(X_1, \ldots, X_n)\, d\mu = \int h(Z_1, \ldots, Z_n)\, d\mathbb{P}.$$

Recall we are no longer in Dubins Savage framework but an arbitrary finitely additive probability space. We have started with an arbitrary sequence of $\mathbb{R}$ valued random variables $\{X_n\}$ and we have constructed $\{Z_n\}$ on a counably additive probability space such that for all n, for all $h \in C_b(\mathbb{R}^n)$

$$\int h(X_1, \ldots, X_n)\, d\mu = \int h(Z_1, \ldots, Z_n)\, d\mathbb{P}.$$

We have not assumed any independence.

Essentially the same construction as before yields here too $\{Y_n\}$ and $\{W_n\}$ such that $(Y_1, \ldots, Y_m)$ and $(Z_1, Z_2, \ldots, Z_m)$ have the same law for all m and

$$\mu^*(|X_n - Y_n| \geq \frac{1}{2^n}) \leq \frac{1}{2^n}$$

$$\mathbb{P}(|Z_n - W_n| \geq \frac{1}{2^n}) \leq \frac{1}{2^n}$$

Thus the same arguments as in the earlier case yield:

Let $g_n : \mathbb{R}^\infty \mapsto S$ be a sequence of functions such that

$$|g_n(x_1, x_2, \ldots) - g_n(y_1, y_2, \ldots)| \sum_{j=1}^{\infty} C_{n,j} |x_j - y_j|^p$$

for $0 < p < \infty$, and $c_{n,j}$ are constants such that $0 \leq C_{n,j} \leq 1$ and $\lim_n C_{n,j} = 0$ for each j.

Theorem. Let λ be a counably additive probability measure on S and let $\alpha \in S$.

Then

$$g_n(X_1, X_2, \dots) \xrightarrow{d} \lambda \text{ if and only if } g_n(Z_1, Z_2, \dots) \xrightarrow{d} \lambda$$

and

$$g_n(X_1, X_2, \dots) \xrightarrow{\sigma} \alpha \text{ if and only if } g_n(Z_1, Z_2, \dots) \xrightarrow{\mathbb{P}} \alpha$$

When it comes to almost sure type limit theorems, we have already seen that even SLLN need not be valid.

We will say that a sequence $\{V_n\}$ defined on a fintely additive probability space is *regular* if for a sequence of functions $\{\psi_n\}$ such that each ψ_n takes finitely many values, the law of the sequence $\{\psi_n(V_n)\}$ is countably additive.

The coordinate random variables in the Dubins Savage strategic framework are regular.

If the sequecne $\{X_n\}$ is regular, then for any limit statute A

$$(X_1, X_2, \ldots) \in A \; a.s. \text{ if and only if } (Z_1, Z_2, \ldots) \in A \; a.s.$$

For each n let E_n be the set of continuity points of $t \mapsto \mu(X_n \leq t)$. It can be shown that for $a_k, b_k \in E_k$, $1 \leq k \leq n$

$$\int h(X_1, \ldots, X_n) \prod_{k=1}^{n} \mathbf{1}_{\{X_k \in [a_k, b_k]\}} \, d\mu = \int h(Z_1, \ldots, Z_n) \prod_{k=1}^{n} \mathbf{1}_{\{Z_k \in [a_k, b_k]\}} \, d\mathbb{P}.$$

Now we can define various dependency notions for the sequence $\{X_n\}$ and they translate to same notion on $\{Z_n\}$.

Mixing: for $1 \leq m \leq n$ let $\mathscr{D}_n^m$ be the field on H generated by the sets

$$\{X_k \leq a_k\},\ a_k \in E_k,\ m \leq k \leq n.$$

We say that $\{X_n\}$ is strongly mixing with rate $r(n)$ if

$$|\mu(A \cap B) - \mu(A)\mu(B)| \leq r(n),$$

for all $A \in \mathscr{D}_k^1$, $B \in \mathscr{D}_m^{k+n}$ for all $k \leq k+n \leq m$.
Now it is clear that if $\{X_n\}$ is strongly mixing with rte $r(n)$ then so is $\{Z_n\}$ with same rate function.

Say that $\{X_n\}$ is a mean zero weakly stationary sequence if

$$\int X_n \, d\mu = 0, \int X_n X_{n+k} \, d\mu = \int X_1 X_{1+k} \, d\mu.$$

Once again it is clear that $\{X_n\}$ is a mean zero weakly stationary sequence if and only if $\{Z_n\}$ is a mean zero weakly stationary sequence.

Say that $\{X_n\}$ is a martingale if

$$\int X_m \mathbf{1}_A \, d\mu = \int X_n \mathbf{1}_A \, d\mu]$$

for all $A \in \mathscr{D}_n^1$, $n \leq m$. Once again it is clear that $\{X_n\}$ is a Martingale if and only if $\{Z_n\}$ is a martingale.

Likewise we can define strictly stationary, ϕ- mixing, β-mixing etc for $\{X_n\}$ and it translates to the same notion on $\{Z_n\}$ in the usual framwork.

Thus all limit laws of the type convergence in probability and convergence in distribution that are valid for $Z_n\}$ continue to hold for $\{X_n\}$ under same conditions- which are a combination of conditions on moments and dependence.

Further if $\{X_n\}$ is regular the same is true for almost sure limit theorems.