BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Measure and Probability Theory

Course Code: 23ST02CC

Unit-II

Measure Theory

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

Ms. S. Soundarya
Guest Faculty
Department of Statistics

UNIT – II

LEBESGUE MEASURE

Lebesgue Measure

In this section we shall define Lebesgue Measure, which is a generalization of the idea of length.

Definition: The length (I) of an interval I with end points a and b is defined as the difference of the end points. In symbols, we write.

$$l\,(I) = b - a.$$

Definition: A function whose domain of definition is a class of sets is called a Set Function. For example, length is a set function. The domain being the collection of all intervals.

Definition. An extended real – valued set function μ defined on a class E of sets is called Additive if

$$A \in E, B \in E, A \cup B \in E \text{ and } A \cap B = \phi, imply$$

$$\mu\,(A \cup B) = \mu\,(A) + \mu\,(B)$$

Definition: An extended real valued set function μ defined on a class E of sets is called finitely additive if for every finite disjoint classes $\{A1, A2, \dots, An\}$of sets in E, whose union is also in E,

we have

$$\mu(U_{i=1}^{n} A_i\,) = \sum_{i=1}^{n} \mu(A_i)$$

Definition. An extended real–valued set function μ defined on a class E of sets is called countably additive it for every disjoint sequence $\{An\}$ of sets in E whose union is also in E, we have

$$\mu(U_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

Definition. Length of an open set is defined to be the sum of lengths of the open intervals of which it is composed of. Thus, if G is an open set, then

$$l(G) = \sum_{n} l(I_n)$$

Where

$$G = U_n I_n, I_{n_1} \cap I_{n_2} = \phi \text{ if } n_1 \neq n_2.$$

Definition. The Lebesgue Outer Measure or simply the outer measure m* of a set A is defined as

$$m^*(A) = \inf_{A \subseteq U I_n} \sum l(I_n).$$

where the infimum is taken over all finite or countable collections of intervals $\{I_n\}$ such that

$$A \subseteq U I_n$$

Since the lengths are positive numbers, it follows from the definition of $m*$ that $m*(A) \geq 0$.

Lebesgue-Stieltjes measure

A cumulative distribution function, or distribution function for short, is a right-continuous, nondecreasing function F: R → R such that

$$\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to \infty} F(x) = 1.$$

If (Ω, F, P) is a probability space and X is a real random variable defined on Ω then the cumulative distribution function (c.d.f.) of X is the distribution function

$$F_x(x) = P(X \leq x).$$

That $F_x(x)$ is in fact a distribution function follows from the upper and lower continuity of probability measures.

If $\mu=P$ is a probability measure then is continuous from above and below, i.e.,

$$A1 \subset A2 \subset A3 \subset \ldots \Rightarrow P(\bigcup_{n\geq 1} An) = \text{limn}\to\infty\ P(An),$$

$$A1 \supset A2 \supset A3 \supset \ldots \Rightarrow P(\bigcap_{n\geq 1} An) = \text{limn}\to\infty\ P(An)$$

Measure integral

1. If f is a nonnegative simple function, so that $f = \Sigma_{i \in I}\, a_i 1_{A_i}$ where I is a finite index set, $a_i \in [0, \infty)$ for i ∈ I, and $\{A_i : i \in I\}$ is measurable partition of S, then

$$\int_S f\, d\mu = \sum_{i\in I} a_i\, \mu(A_i)$$

2. If f: S→ [0,∞) is measurable, then

$$\int_s f\ du = \{\int_s g du : g \text{ is simple and } 0 \leq g \leq 1\}\ f\ S$$

3. 3 . If f: S→ ℝ is measurable, then

$$\int_s f\, d\mu = \int_s f^+\, d\mu - \int_s f^-\, d\mu$$

4. If f: S→ ℝ is measurable and A ∈ S, then the integral of f over A is defined by

$$\int_A f\, d\mu = \int_s 1_A f\, d\mu$$

assuming that the integral on the right exists.

Properties of Measure integral

1. If f, g : S → R are measurable functions whose integrals exist, then $\int_S (f + g)\, d\mu = \int_S f\, d\mu + \int_S g\, d\mu$ as long as the right side is not of the form $\infty - \infty$.
2. If f: S → R is a measurable function whose integral exists and c ∈ R, then $\int_S cf\, d\mu = c \int_S f\, d\mu$.
3. If $f: S \to \mathbb{R}$ is measurable and $f \geq 0$ on S then $\mu \geq 0$.
4. If f,: $S \to \mathbb{R}$ are measurable functions whose integrals exist and $f \leq g$ on S then $\int_S f d\mu \leq \int$

Comparison of Lebesgue and Riemann integration

(1) The most obvious difference is that in Lebesgue's definition we divide up the interval into subsets while in the case of Riemann we divide it into subintervals.

(2) In both Riemann's and Lebesgue's definitions we have upper and lower sums which tend to limits. In Riemann case the two integrals are not necessarily the same and the function is integrable only if they are same. In the Lebesgue case the two integrals are necessarily the same, their equality being consequence of the assumption that the function is measurable.

(3) Lebesgue's definition is more general than Riemann. We know that if function is the R-integrable then it is Lebesgue integrable also, but the converse need not be true. For example the characteristic function of the set of irrational points have Lebesgue integral but is not R-integrable.

Let χ be the characteristic function of the irrational numbers in [0,1]. Let E_1 be the set of irrational number in [0,1], and let E_2 be the set of rational number in [0,1]. Then $P = [E_1, E_2]$ is a measurable partition of (0,1]. Moreover, χ is identically 1 on E_1 and χ is identically 0 on E_2 .

Hence $M[\chi, E_1] = m[\chi, E2] = 1$, while $M[\chi, E_1] = m[\chi, E_2] = 0$. Hence $U[\chi, P] = 1.m\ E_1 + 0.m\ E_2 = 1$. Similarly $L[\chi, P] = 1.m\ E_1 + 0.M\ E_2 = 1$. Therefore, $U[\chi, P] = L[\chi, P]$.

For Riemann integration

$$M[\chi,J] = 1 \ , \ m[\chi,J] = 0$$

for any interval $J\subset[0,1]$

$$\therefore U[\chi,J]=1, L[\chi,J]=0 \ .$$

The function is not Riemann- integrable.

Lebesgue integral

Lebesgue integration is an alternative way of defining the integral in terms of measure theory that is used to inte grate a much broader class of functions than the Riemann integral or even the Riemann Stieltjes integral. The idea behind the Lebesgue integral is that instead of approximating the total area by dividing it into vertical strips, one approximates the total area by dividing it into horizontal strips.

The shortcomings of the Riemann integral suggested the further investigations in the theory of integration. We give a resume of the Riemann Integral first.

Let f be a bounded real- valued function on the interval [a, b] and let

$$a = \xi_0 < \xi_1 < \cdots < \xi_n = b$$

Be a partition of [a, b]. Then for each partition we define the sums

$$S = \sum_{i=1}^{n}(\xi_i - \xi_{i-1})M_i$$

And

$$s = \sum_{i=1}^{n}(\xi_i - \xi_{i-1})m_i$$

Where

$$M_i = \sup_{\xi_{i-1} < x < \xi_i} f(x)\ ,\ m_i = \inf_{\xi_{i-1} < x < \xi_i} f(x)$$

We then define the upper Riemann integral of f by

$$R \overline{\int_a^b} f(x)dx = \inf S$$

With the infimum taken over all possible subdivisions of [a, b].

Similarly, we define the lower integral

$$R \underline{\int_a^b} f(x)dx = \sup s.$$

The upper integral is always at least as large as the lower integral, and if the two are equal we say that f is Riemann integrable and call this common value the Riemann integral of f. We shall denote it by

$$R \int_a^b f(x)$$

To distinguish it from the Lebesgue integral, which we shall consider later. By a step function we mean a function ψ which has the form

$$\psi(x) = c_i\ ,\ \xi_{i-1} < x < \xi_i$$

for some subdivision of [a, b] and some set of constants c_i.

The integral of ψ(x) is defined by

$$R \int_a^b \psi(x)\, dx = \sum_{i=1}^{n} c_i \left(\xi_i - \xi_{i-1}\right).$$

With this in mind we see that

$$R \overline{\int_a^b} f(x)dx = \inf \int_a^b \psi(x)\, dx$$

for all step function

$$\psi(x) \geq f(x).$$

Similarly,

$$R \int_a^b f(x)dx = \sup \int_a^b \phi(x)\, dx$$

for all step functions

$$\phi(x) \leq f(x).$$

Dominated convergence theorem

$(f_n(x))^{\infty}_{n=1}$ denotes a sequence of Lebesgue integrable functions that almost everywhere on I converge to a limit function f. Assume that g is a Lebesgue integrable function that $|f_n| \leq g$ nearly everywhere on I and for all $n \in N$.

If $\lim_{n\to\infty} \int_I f_n(x)\, dx = \int_I f(x)dx.$,

then f is Lebesgue integrable on I.

Note:

It's important to remember that $\lim_{n\to\infty} \int_I f_n(x)\, dx = \int_I f(x)\, dx$ can alternatively be written as $\lim_{n\to\infty} \int_I f_n (x)\, dx = \int_I \lim_{n\to\infty} fn (x)\, dx$.

Proof:

(fn(x))n=1 is a sequence of Lebesgue integrable functions that nearly everywhere on I converge to a limit function f. Then, practically everywhere on I, the following equality holds:

$$\text{Lim sup}_{n\to\infty} f_n(x) = f(x) = \lim \text{inf}_{n\to\infty} f_n(x) \ \ldots(1)$$

two new sequences of functions $(G_n(x))^{\infty}_{n=1}$ and $(g_n(x))^{\infty}_{n=1}$

$$G_n(x) = \sup_{k\geq n} \{f_k(x)\} \text{ and } g_n(x) = \inf_{k\geq n} \{f_k(x)\} \ \ldots(2)$$

$(G_n(x))^\infty_{n=1}$ is a decreasing sequence of functions , $(g_n(x))^\infty_{n=1}$ is an expanding sequence of functions that almost always converges to f(x) on I. $(G_n(x))\infty_{n=1}$ and $(g_n(x))^\infty_{n=1}$ are Lebesgue integrable function sequences.

almost everywhere on I, the following inequality holds:

$$g_n(x) \leq f(x) \leq G_n(x) \quad \ldots(3)$$

Already know that

$$|f_n(x)| \leq g(x)$$

And

some Lebesgue integrable function g. As a result, nearly everywhere on I, the following chain of inequalities

$$-g(x) \leq g_n(x) \leq f_n(x) \leq G_n(x) \leq g(x) \quad \ldots(4)$$

the function sequence $(g-G_n)^\infty_{n=1}$. Then $g-G_n$ is Lebesgue integrable on I and, moreover, is nearly everywhere on I an expanding sequence.

the sequence is

$$\left(\int_I [g(x) - G_n(x)]\, dx\right)^\infty_{n=1}$$

$$\int_I [g(x) - G_n(x)]\, dx \leq \int_I 2g(x)\, dx = 2 \int_I g(x)\, dx < \infty \quad \ldots(5)$$

the function $g(x)-f(x) = \lim_{n\to\infty} [g(x) - Gn(x)]$ is Lebesgue integrable on I using Levi's Monotonic Convergence theorems concerning Lebesgue integrable functions, and that:

$$\int_I [g(x) - f(x)] = \lim_{n\to\infty} \int_I [g(x) - G_n(x)]dx = \int_I g(x)dx - \lim_{n\to\infty} \int_I G_n(x)dx = \int_I g(x)dx - \int_I f(x)dx \quad \ldots(6)$$

As a result, $\int_I f(x)dx = \lim_{n\to\infty} \int_I G_n(x)dx$, and f is Lebesgue integrable on I.

The function sequence

$$(g_n(x) + g(x))^\infty_{n=1}.$$

This is a sequence of Lebesgue integrable functions on I in increasing order. Moreover,

$$\left(\int_I [g_n(x) + g(x)]dx \right)^{\infty}_{n=1},$$

the sequence converges, because

$$\int_I [g_n(x) + g(x)]\, dx \leq \int_I 2g(x)dx = 2\int_I g(x)dx < \infty \ldots(7)$$

So we know that

$$g(x) + f(x) = \lim_{n\to\infty} [g(x) + g_n(x)]$$

is a Lebesgue integrable function according to Levi's Convergence Theorem regarding Lebesgue integrable functions, that:

$$\int_I [g(x) + f(x)]\, dx = \lim_{n\to\infty} \int_I [g(x) + g_n(x)]\, dx = \int_I g(x)dx + \lim_{n\to\infty} \int_I g_n(x)dx = \int_I g(x)dx + \int_I f(x)dx \ldots (8)$$

is Lebesgue integrable on I, and that $\int_I f(x)dx = \lim_{n\to\infty} \int_I g_n(x)\, dx$ once more.

$g_n(x) \leq f(x) \leq G_n(x)$ holds almost everywhere on I, the preceding chain of inequalities holds nearly everywhere on I:

$$\int_I g_n(x)dx \leq \int_I f(x)dx \leq \int_I G_n(x)dx \lim_{n\to\infty} \int_I g_n(x)dx \leq \lim_{n\to\infty} \int_I f(x)dx \leq \lim_{n\to\infty} \int_I G_n(x)dx \int_I f(x)dx \leq \lim_{n\to\infty} \int_I f(x)dx \leq \int_I f(x)dx \ldots(9)$$

Therefore,

$$\lim_{n\to\infty} \int_I f_n(x)dx = \int_I f(x)dx \text{ is proved.}$$

Radon–Nikodymn theorem

Statement:

Let (X, M, μ) be a σ - finite measure space and v a σ - finite measure defined on the measurable space (X, M) that is absolutely continuous with respect to μ. Then there is a non-negative function f on X that is measurable with respect to M, for which

$$v(E) = \int_E f\, d\mu,\ \forall\ E \in M \ ....(1)$$

The function f is unique in the sense that if g is any non-negative measurable function on X that also has this property, then

$$g = f \text{ a.e } [\mu].$$

Proof:

μ and v finite measures and then we extend the result to the finite case.

Assume that,

σ - finite.

If $v(E) = 0,\ \forall\ E \in \mu$ the claim holds with $f = 0$ on X.

Claim:

First we prove that there is non-negative measurable function f on x for which,

$$\int f\, d\mu > 0 \text{ and } \int f\, d\mu \geq v(F)\ \forall\ E \in \mu\ ......(2)$$

From $\lambda > 0$

Hahn Decomposition theorem,$\{ P_\lambda , N_\lambda \}$ for $v - \lambda\mu$.

Where,

$X = P_\lambda \cup N_\lambda,\ P_\lambda \cap N_\lambda = \Phi$.

$P\lambda$ is a positive set and $N\lambda$ is a negative set with respect to the signed measure $v - \lambda\mu$.

Claim: there is some $\lambda > 0$ for which $\mu(P\lambda) > 0$.

Assume,

$\mu\ (P_\lambda) = 0$ for all $\lambda > 0$.

Then for any measurable subset E of P_λ, we have $\mu(E) = 0$ and hence $V(E) = 0$ by absolute continuity.

Since, N_λ is a negative set for $(V - \lambda\mu)$ for any $E \in \mu$

$(V - \lambda\mu)(E) = (V - \lambda\mu)((E \cap P_\lambda) \cup (E \cap N_\lambda))$

$= (V - \lambda\mu)((E \cap P\lambda) \cup (E \cap N_\lambda))$

$= V((E \cap P_\lambda) \cup (E \cap N_\lambda)) - \lambda\mu((E \cap P_\lambda) \cup (E \cap N_\lambda))$

$= V((E \cap P_\lambda)) + V((E \cap N_\lambda)) - \lambda\mu((E \cap P_\lambda))$

$= V((E \cap N_\lambda)) - \lambda\mu((E \cap N_\lambda))$

$\leq V((E \cap N_\lambda))$

$\leq V(E) - \lambda\mu(E) \leq 0.$

By definition, $V(E) - \lambda\mu(E) \leq 0$ $V(E) \leq \lambda\mu(E)$ for all $E \in \mu$ and all $\lambda > 0$ ……..(3)

Since μ is a finite measure,

$\mu (E) \leq \mu(X) < \infty$ by condition.

$V(E) = 0 \ \forall \ E \in M.$

V does not vanish on all of M, so, the assumption that

$\mu (P_\lambda) = 0 \ \forall \ \lambda > 0$ is false and there is a $\lambda_o > 0$ for which $\mu (P_{\lambda o}) > 0$

Define $f = \lambda_o \chi P_{\lambda o}$ then,

$\int f \, d\mu = \int \lambda_o \chi P_{\lambda o} \, d\mu = \lambda_o \, \mu(P_{\lambda o}) > 0$ and,

$\int f \, d\mu = \int \lambda_o \chi P_{\lambda o} \, d\mu$

$= \int \lambda_o \chi P_{\lambda o} X E \, d\mu$

$= \int \lambda_o \chi (P_{\lambda o} \cap E) \, d\mu$

$= \lambda_o \, \mu (P_{\lambda o} \cap E)$

$\leq V (P_{\lambda o} \cap E)$

$\leq V (P_{\lambda o} \cap F)$

$\leq \int f \, d\mu$ (F) for all $E \in M$

Hence, (2) holds for $f = \lambda_o \chi P_{\lambda o}$

Define F to be collection of non-negative measurable functions on X for which

$\int f \, d\mu \geq v(E)$ for all $E \in \mu$.

Then, Fμ is non-empty since,

$\lambda_o \chi P_{\lambda o} \in F$

$M=\sup_{f\in F} \int f \, d\mu$ ....(4)

NOTE: $M > 0$, since $\lambda_o \chi P_{\lambda o} \in F$

claim: There is an $f \in F$ for which $\int_\lambda f \, d\mu = M$, for any such f,

$$v(E) = \int_E f \, d\mu \text{ for all } E \in M$$

let, $g, h \in F$ and $F \in M$

Then, $E_1 = \{ x \in E \mid g(x) \leq h(x) \}$

We have, $E_2 = \{ x \in E \mid g(x) \geq h(x) \}$

$\int \max \{g, h\} \, d\mu = \int_{E1} h \, d\mu + \int_{E2} g \, d\mu$

$\geq v(E_1) + v(E_2)$ [by def. of F]

$\geq v(E)$

So, that $\max \{g, h\} \in F$.

Select, a sequence $\{f_n\}$ in F for which $\lim \int f_n \, d\mu = M$ (such a sequence exists by def. of supremum)

Assume, $\{f_n\}$ to be a pointwise increasing sequence of function of x [or else we can replace each f_n by maximum $\{f_1, f_2, f_3, ..., f_n\} \in F$ to get an increasing sequence]

Define, $f(x) = \lim f_n(x)$ for each $x \in X$.

Since F consists only non-negative functions and $\{f_n\}$ is monotone, by monotone convergence theorem

$\int_X f\, d\mu = \lim \int_X f_n\, d\mu = M$.

Also, $v(E) \geq \int_E f_n d\mu\ \forall\ E \in M$ and as by the monotone convergence theorem

$V(E) \geq \lim \int_E f_n\, d\mu = \int \lim f\, d\mu\ \forall\ E \in M$.

So, $f \in F$.

Define $\eta\ (F) = V(E)\ .\ \int_E f\, d\mu$ for $E \in M$

By assuming $V(X) < \infty$. Therefore, $\int_X f\, d\mu \leq V(X) < \infty$.

Since, $V(E) \geq \int_E f\, d\mu$ for all

$E \in M$, $\eta\ (E) \geq 0$ for all $E \in M$

Now, ν is countably additive since it is a measure and so h is countably additive. Therefore η is a measure on M.

Also, for $E \in M$ with $\mu(E) = 0$ we have $\nu(E) = 0$ since ν is absolutely continuous with respect to μ and $\int_E f\, d\mu = 0$. So that, $\eta(E) = 0$ that is η is absolutely continuous with respect to μ.

We claim that $\eta = 0$ on M.

Assume there is some set $E \in M$ for which $\eta(E) > 0$. Then, as argued above v, we can find a non-negative function f such that $\int \hat{f}\, d\mu > 0$ and $\int \hat{f} d\mu \leq \eta\ (E)$ for all $E \in M$.

So, from the definition of η we have for this $\hat{f}$ that, $\int_E \hat{f}\, d\mu \leq \eta(E) = \nu(E) - \int_E \hat{f}\, d\mu$ for all $E \in M$.

Then,

$$\int(f+\tilde{f})\, d\mu=\int f\, d\mu+\int \tilde{f}\, d\mu >0$$

and for all $E \in M$

$$\int_E (f+\tilde{f})d\mu = \int_E f\, d\mu+\int_E \tilde{f}\, d\mu \leq \nu(E)$$

so that $f+\tilde{f} \in F$

But, then

$$\int_X (f+\tilde{f})\, d\mu > \int_X \tilde{f} d\mu=M$$

a Contradiction of the definition of M. So, that assumption that η (E)>0 for some set

E∈M is false and hence η (E)=0 ∀ E∈ M.

$$\nu(F)=\int f d\mu \ \forall E\in M$$

Almost everywhere convergence

A Sequence {fn} of functions defined on a set E is said to converge almost everywhere to f if

$$\lim_{n\to\infty} f_n(x) = f(x) \ \forall x \in E - E_1 \ where \ E_1 \subset E,$$

$$m(E_1) = 0.$$

Convergence in measure

Definition: A sequence < fn > of measurable functions is said to convergence in measure to f on a set E,

$$m\{x||f(x) - f_n(x)| \geq \varepsilon\} < \delta.$$

Or

$$\lim_{n\to\infty} m\{x||f(x) - f_n(x)| \geq \varepsilon\} = 0$$

Convergence in mean

Let r≥1 be a fixed number. A sequence of random variables X1, X2, X3, ….. converges in the r^{th} mean to a random variable X, shown by

$$X_n \xrightarrow{L^r} X, \text{ if}$$

$$\lim_{n\to\infty} E\left(|X_n - X|^r\right) = 0.$$

If r=2, it is called the mean-square convergence, and it is shown by

$$X_n \xrightarrow{m.s.} X.$$