Lebesgue Integrals

Tuesday, October 11, 2011
12:32 PM

Sponsored by the National Grid Foundation nationalgrid The power of action.

① Integral of a measurable function with respect to a prob. measure. [Definition, Basic properties]

② Three main convergence theorems
[Monotone Convergence Th, Fatou's Lemma, Lebesgue Dominated Convergence Th.]

③ "Change of variable" formula

④ Radon-Nikodym theorem.

---

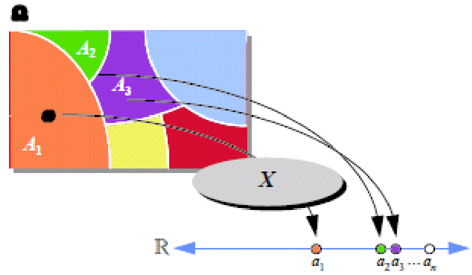
① Let $(\Omega, \mathcal{F}, P)$ be a probability space

Definition 1: A random variable $X : \Omega \to \mathbb{R}$ is called simple if it takes at most countable number of values

Examples:

1) $A \in \mathcal{F}$: let $X(\omega) = \mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$

2) $A_1, A_2, \ldots, A_n \in \mathcal{F}$; $A_i \cap A_j = \emptyset$ (for $i \neq j$)
set $X(\omega) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega)$
(for some constants $a_1, a_2, \ldots, a_n$)



3) $A_1, A_2, \ldots \in \mathcal{F}$; $A_i \cap A_j = \phi$ (for $i \neq j$) and $\{a_n\}_{n \geq 1}$

such that $a_i \geq 0$, $\sum_{i=1}^{\infty}{}' a_i < \infty$;

set $X(\omega) = \sum_{i=1}^{\infty}{}' a_i 1_{A_i}(\omega)$

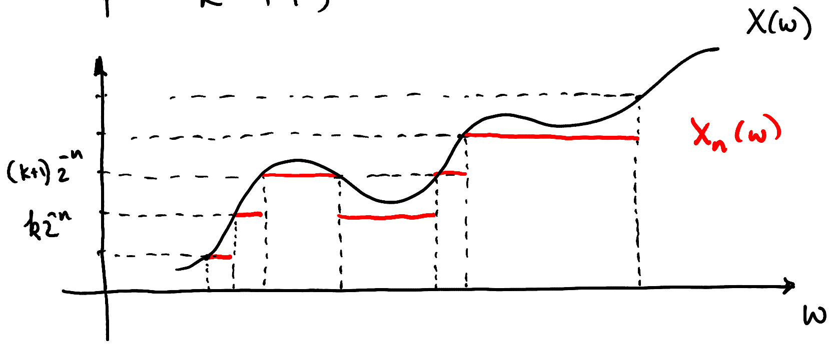
Remark: sums, products, quotients of simple functions are also simple

Theorem 1 : Let X be a non-negative random variable.
There is a sequence of nonnegative simple random variables $\{X_n\}$ such that $(X_n \uparrow X)$

(1) $X_n(\omega) \leq X_{n+1}(\omega)$ for all $\omega \in \Omega$

(2) $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ for all $\omega \in \Omega$

Proof: Define $X_n(\omega) = k 2^{-n}$ if $k 2^{-n} \leq X(\omega) < (k+1) 2^{-n}$
for $k = 0, 1, 2, \ldots$



It is easy to see that (1) & (2) are satisfied.

It is easy to see that (1) & (2) are satisfied.

The integral is defined in several steps:

(I) for non negative simple random variables

(II) for general non negative random variables

(III) for general random variables (may not exist all the time)

---

let $(\Omega, \mathcal{F}, P)$: be a given prob. space.

(I) let $X \geq 0$ be a simple random variable (countable)

$X \in \{a_1, a_2, \dots\}$

$A_i = \{\omega \in \Omega : X(\omega) = a_i\} = X^{-1}(\{a_i\})$

Since X is measurable, $A_i \in \mathcal{F}$ for all i.

$$EX = \int_\Omega X(\omega)\, dP(\omega) := \sum_{i=1}^{\infty} a_i P(A_i)$$

Remark: this EX can be ∞

Properties:

(1) $EX \geq 0$

(2) $E\mathbb{1}_A = P(A)$ for every $A \in \mathcal{F}$

(3) $E(\alpha X + \beta Y) = \alpha EX + \beta EY$ for all $\alpha, \beta > 0$

(4) If $X \geq Y \geq 0$ then $EX \geq EY$

(II) let $X \geq 0$ be a random variable

Define $EX = \int_\Omega X(\omega)\, dP(\omega) = \lim_{n \to \infty} EX_n$

where $\{X_n\}$ is a sequence of non-negative simple r.v

such that $X_n \uparrow X$ (as $n \to \infty$)

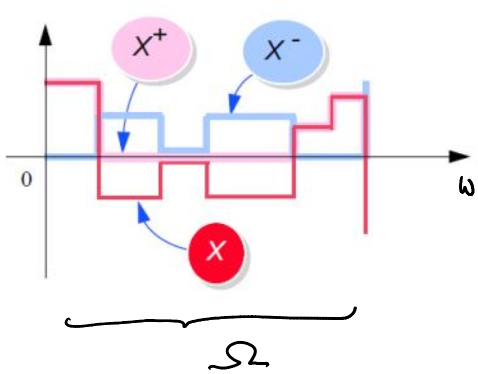
Theorem 2: EX does not depend on the approximating sequence

Proof: (Koralov, Sinai, p 38-39)

III) Let X be an arbitrary random variable

$\rightsquigarrow$ define $X_+(\omega) := X(\omega) \vee 0$; $X_-(\omega) := (-X(\omega)) \vee 0$

$\Rightarrow$ $X_+, X_- \geq 0$, also $X(\omega) = X_+(\omega) - X_-(\omega)$



Remark:

$$|X(\omega)| = X_+(\omega) + X_-(\omega)$$

- by II EX_+ and EX_- are well defined.
- if at least one of them is finite we define:

$$EX = EX_+ - EX_-$$

(with the convention that $\infty - c = \infty$; $c - \infty = -\infty$)

- when both EX_+ and EX_- are infinite $\Rightarrow$ EX is undefined

Theorem 3: If EX or EY is finite and $\alpha, \beta \in \mathbb{R}$

- $E(\alpha X + \beta Y) = \alpha EX + \beta EY$
- if $X \geq Y$ then $EX \geq EY$

Definition 2: A random variable X is said to be integrable (written as $X \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$, $\mathcal{F}, P$ are often dropped from the notation) if $EX_+ < \infty$ and $EX_- < \infty$

Remark:

1) $|X| = X_+ + X_- \Rightarrow X \in \mathcal{L}^1(\Omega)$ iff $|X| \in \mathcal{L}^1(\Omega)$

2) If $A \in \mathcal{F}$ and X is measurable then $X 1_A$ is also measurable and

$$\int_A X(\omega)\, dP(\omega) = \int_\Omega X(\omega) 1_A(\omega)\, dP(\omega) = E[X 1_A]$$

Theorem 4 (σ-additivity of the Lebesgue integral)

Let $A = \bigcup_{i=1}^{\infty} A_i$ where $A_i \in \mathcal{F}$, $A_i \cap A_j = \phi$ $(i \neq j)$ and X is a random variable such that $E|X 1_A| < \infty$. Then

$$E X 1_A = \int_A X(\omega)\, dP(\omega) = \sum_{i=1}^{\infty} \int_{A_i} X(\omega)\, dP(\omega) = \sum_{i=1}^{\infty} E(X 1_{A_i})$$

$\rightsquigarrow$ Theorems 3 & 4 give the most basic properties of the Lebesgue integral.

$\rightsquigarrow$ 3 more useful properties need to be mentioned as well

Definition 3: We say that X is square integrable or $X \in \mathcal{L}^2(\Omega)$ if $X^2 \in \mathcal{L}^1(\Omega)$

Markov's inequality: If $X \geq 0 \Rightarrow P(X \geq a) \leq \frac{EX}{a}$ $(\forall) a$.

Chebyshev's inequality: $P(|X| \geq a) \leq \frac{EX^2}{a^2}$

- $P(|X - EX| \geq a) \leq \frac{\operatorname{Var} X}{a^2}$

Cauchy-Schwarz inequality: $X, Y \in \mathcal{L}^2(\Omega) \Rightarrow XY \in \mathcal{L}^1(\Omega)$ and

$$|E(XY)| \leq E|XY| \leq (EX^2)^{1/2} (EY^2)^{1/2}$$

Proof: $|XY| \leq \frac{1}{2}(X^2 + Y^2) \Rightarrow XY \in \mathcal{L}^1(\Omega)$

- also easy to see $EXY \leq E|XY|$ and $E(-XY) \leq E|XY|$

so $|E(XY)| \leq E|XY|$

- let $f(t) = E(t|X| + |Y|)^2 \Rightarrow f(t) \geq 0$, well defined

$$f(t) = t^2 EX^2 + 2t E|XY| + EY^2 \geq 0 \quad (\forall) t$$

$\Rightarrow$ the discriminant $4[E|XY|]^2 - 4EX^2EY^2 \leq 0 \quad (\forall) t$

$$\Rightarrow E|XY| \leq (EX^2)^{1/2} (EY^2)^{1/2}$$

Jensen inequality: $\varphi: \mathbb{R} \to \mathbb{R}$ be a convex function and X is a random variable such that $E|\varphi(X)| < \infty$. Then

$$\varphi(EX) \leq E\varphi(X)$$

Definition $\varphi: \mathbb{R} \to \mathbb{R}$ is said to be **convex** if for all $x, y \in \mathbb{R}$, and $\lambda \in [0,1]$

$$\varphi(\lambda x + (1-\lambda) y) \leq \lambda \varphi(x) + (1-\lambda)\varphi(y)$$

- alternative definition of a **convex** function is

$\varphi(x) = \sup_{n} (a_n x + b_n)$ for a countable collection of pairs of real numbers (a_n, b_n)

Examples :

- $|EX| \le E|X|$: $\varphi(x) = |x|$
- $(E|X|)^p \le E|X|^p$; $p \ge 1$: $\varphi(x) = |x|^p$
- $e^{EX} \le E e^X$; $\varphi(x) = e^x$

② Three main convergence theorems

$X_n \to X$ a.s. means : $P(\{\omega : X_n(\omega) \not\to X(\omega)\}) = 0$

(a) Monotone convergence theorem : $\{X_n\}_{n \ge 1}$; $X_n \ge 0$ a.s. and $X_n \uparrow X$ a.s. then $\boxed{\lim_{n\to\infty} EX_n = EX}$

$(X_n \le X_{n+1})$

(b) Fatou's lemma : $\{X_n\}_{n\ge1}$; $X_n \ge 0$ a.s. then

$$\boxed{E\left[\liminf_{n\to\infty} X_n\right] \le \liminf_{n\to\infty} EX_n}$$

(c) Lebesgue dominated convergence theorem :

If $X_n \to X$ a.s. and $|X_n| \le Y$ a.s for all n, where $Y \in \mathcal{L}^1(\Omega)$

Then $X_n \in \mathcal{L}^1(\Omega)$, $X \in \mathcal{L}^1(\Omega)$ and

$$\boxed{\lim_{n\to\infty} EX_n = EX} \qquad \lim_{n\to\infty} \int_\Omega X_n \, dP = \int_\Omega \overbrace{\lim_{n\to\infty} X_n}^{X} \, dP$$

Example: $(\Omega, \mathcal{F}, P) = \left([0,1], \mathcal{B}_{[0,1]}, \text{uniform measure}\right)$

$$X_n(\omega) = \begin{cases} n & \text{if } \omega \in [0, \frac{1}{n}] \\ 0 & \text{if } \omega \in (\frac{1}{n}, 1] \end{cases}$$

· clearly $X_n \xrightarrow[n\to\infty]{} 0$ for all ω except $\omega = 0$!

$EX_n = n \cdot \frac{1}{n} = 1 > 0 = E\, 0$

(strict inequality in Fatou's lemma)

Change of variable formula:

Let $(E, \mathcal{E})$ be a measurable space, X measurable function from Ω to E, and h a measurable function from E to $\mathbb{R}$.

Then:

(a) $h(X) \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ iff $h \in \mathcal{L}^1(E, \mathcal{E}, P_X)$

(b) If $h \geq 0$ or $h(X) \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ then

$$Eh(X) = \int_\Omega h(X(\omega))\, dP(\omega) = \int_E h(x)\, dP_X(x)$$

Recall: $\boxed{P_X(A) = P(X^{-1}(A))}$ for every $A \in \mathcal{B}$

Example: Let $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$ and X have density $f(x)$ and $P_X((a,b]) = \int_a^b f(x)\, dx$. Then

$$Eh(X) = \int_{-\infty}^{+\infty} h(x) \underbrace{dF(x)}_{= dP_X(x)} = \int_{-\infty}^{+\infty} h(x) f(x)\, dx$$

Example : $(E, \mathcal{E}) = (\mathbb{R}^n, \mathcal{B}^n)$ and $X = (X_1, \ldots, X_n)$ have density $f(x)$ i.e

$$P_X\left((a_1,b_1] \times (a_2,b_2] \times \cdots \times (a_n,b_n]\right) = \int_{a_1}^{b_1}\int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1,\ldots,x_n)\, dx_n\, dx_{n-1} \cdots dx_1$$

Then

$$Eh(X) = \int_{\mathbb{R}^n} \cdots \int h(x_1,\ldots,x_n)\, f(x_1,\ldots,x_n)\, dx_1 \cdots dx_n$$

Example : If X is simple and takes values $\{a_1, a_2, \ldots\}$ then $h(X)$ is also simple and

$$Eh(X) = \sum_{i=1}^{\infty} h(a_i)\, P(X = a_i)$$

· notice that the right hand side is $\int_{\mathbb{R}} h(x)\, dP_X(x)$

where P_X is the distribution of X.

· we often write such measure as :

$$P_X(\cdot) = \sum_{i=1}^{\infty} P(X = a_i) \cdot \delta_{a_i}(\cdot) \qquad \text{where}$$

$$\delta_{a_i}(B) = \begin{cases} 1 & \text{if } a_i \in B \\ 0 & \text{otherwise} \end{cases} \qquad \text{"Dirac measure"}$$

Comparison between Riemann & Lebesgue integrals

Definition 6 : **Lebesgue measure** : m on $(\mathbb{R}, \mathcal{B})$ is a function from $\mathcal{B} \to [0,\infty]$ such that function from $\mathcal{B} \to [0,\infty]$ such that

(1) m is countably additive

(2) $m\big((a,b]\big) = b-a$ for every $a,b \in \mathbb{R}$ $a \le b$.

Remark: one needs to prove that such a measure exists and it is unique, but we shall take it for granted.

- given a measurable function $f:\mathbb{R} \to \mathbb{R}$ we can follow the same steps as above to construct its integral with respect to the Lebesgue measure

Theorem 5: Let f be a bounded function on $\mathbb{R}$ and $a<b$.

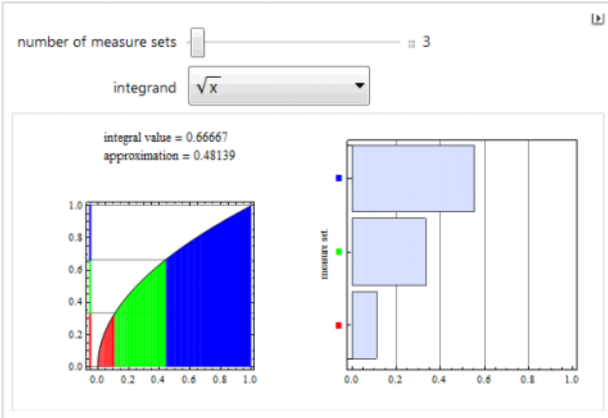
(i) the Riemann integral $\int_a^b f(x)\,dx$ exists iff the Lebesgue measure of the set of discontinuity points of f on $[a,b]$ is equal to 0

(i.e if f is continuous almost everywhere (a.e))

(ii) If $\int_a^b f(x)\,dx$ is defined then f is Borel measurable (i.e Lebesgue integral is also defined as f is bounded) and both integrals agree.

[Lebesgue Integration (Wolfram demonstration)]

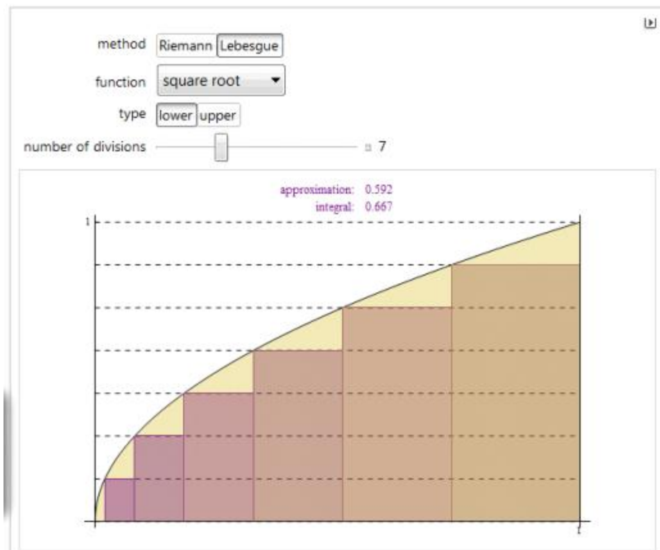
Lebesgue Integration



Lebesgue integration extends the definition of integral to a much larger class of functions than the class of Riemann integrable functions. The Riemann integral is constructed by partitioning the integrand's domain (on the x axis). The Lebesgue integral is constructed by partitioning the integrand's co-domain (on the y axis). For each value y in the co-domain, find the measure $m(y)$ of the corresponding set of points $f^{-1}(y)$ in the domain. Roughly speaking, the Lebesgue integral is then the sum of all the products $y\, m(y)$.

[Riemann vs. Lebesgue Integrals (Wolfram demonstration)]

Riemann versus Lebesgue



The main difference between the Lebesgue and Riemann integrals is that the Lebesgue method takes into account the values of the function, subdividing its range instead of just subdividing the interval on which the function is defined. This fact makes a difference when the function has big oscillations or discontinuities. However, the Lebesgue method needs to compute the measure of sets that are not intervals.

Notation :

Lebesgue integral : $\int_{[a,b]} f(x)\, dm(x)$

Riemann integral : $\int^{b} f(x)\, dx$

Riemann integral : $\int_a^b f(x)\,dx$

Example : $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$

- f is discontinuous at every point $\Rightarrow$ not Riemann integrable (upper Riemann sum is 1 ; lower Riemann sum is 0)

- the Lebesgue integral :

$$\int_{\mathbb{R}} f(x)\,dm(x) = 0 \quad \text{since} \quad m(\{x : f(x) = 1\}) = 0$$

$$= \int_{\mathbb{Q}} dm(x) = m(\mathbb{Q})$$

(4) Change of measure :

$(\Omega, \mathcal{F}, P)$, $(\Omega, \mathcal{F}, \tilde{P})$

- Recall . $P \ll \tilde{P}$ if $\tilde{P}(A) = 0 \Rightarrow P(A) = 0$

- $P \sim \tilde{P}$ if $P \ll \tilde{P}$ and $\tilde{P} \ll P$

  (in other words they agree which sets have measure 0)

Theorem 6 : Let $(\Omega, \mathcal{F}, P)$ be a probability space and $Z \geq 0$ a.s. measurable, and $\mathbb{E}Z = 1$. For $A \in \mathcal{F}$ we define

$$\boxed{\tilde{P}(A) := \int_A Z(\omega)\,dP(\omega)} \rightsquigarrow d\tilde{P}(\omega) = Z(\omega)\,dP(\omega)$$

$$\Downarrow$$

$$\tilde{P} \ll P$$

Then $\tilde{P}$ is a probability measure. Moreover, if $X \geq 0$ and is measurable then

$$\tilde{E} X = E[ZX]$$

- if $Z > 0$ a.s. then $EY = \tilde{E}\left[\frac{1}{Z} \cdot Y\right]$ for $Y \geq 0$. measurable

$$\Downarrow$$

$$P \sim \tilde{P}$$

Proof: Shreve, vol. 2, p 33-34. $\tilde{P}(\Omega) = \int_\Omega Z(\omega) dP(\omega) = 1$

- one needs to check only countable additivity

$\rightsquigarrow$ straightforward consequence of the monotone convergence th.

Remark: The Lebesgue integral $\int_\Omega X \, dP$ is not sensitive to changes of X on a set of measure 0.

More precisely, if $X = Y$ a.s. then $\int_\Omega X \, dP = \int_\Omega Y \, dP$ in case they are well defined.

Radon - Nikodym Theorem: If $\tilde{P} \ll P$ then there is a random variable $Z \geq 0$, $EZ = 1$ such that $(\forall)$ $A \in \mathcal{F}$

$$\tilde{P}(A) = \int_\Omega Z(\omega) \, dP(\omega)$$

$\rightsquigarrow$ this theorem says that the only way to obtain $\tilde{P} \ll P$ is by the procedure described in Theorem 6.