On the Lebesgue measure

Using the Riemann Integral the probability measure in the interval $\{x \le X \le x + dx\}$ is given by

$$P_r\{x \le X \le x + dx\} = f_x(x)\,dx \,,$$ where

$f_x(x)$ is the probability density function

The probability that a R.V. takes on values in the interval $(a, b]$ is given by

$$P_r\{a \le X \le b\} = \int_a^b f_x(x)\,dx = F(b) - F(a)$$

The expectation operation on $g(X)$ is defined via:

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_x(x)\,dx$$

This integral representation is however incomplete in the sense that it requires the integrand $f(\cdot)$ to be bounded & continuous, which may not be the case Hence the need for a more general measure

Specifically assume X is a non-negative random variable

The main idea behind the Lebesgue measure is to divide the interval $(a, b]$ into smaller disjoint intervals of the form $(a_i, b_i]$, $i = 1, 2, 3, \ldots n$

The Lebesgue measure μ associated on the interval $(a, b]$ is :

$$\mu((a, b]) = (b - a).$$

In order for this to be a valid probability measure: $\mu((a, b]) = F(b) - F(a)$
If we divide $(0, \infty]$ into smaller intervals of the form

$$A_i = \left\{ \frac{i}{2^n} \leq X \leq \frac{i+1}{2^n} \right\}$$

The measure associated with each A_i is

$$\mu(A_i) = \Pr\left\{ \frac{i}{2^n} \leq X \leq \frac{i+1}{2^n} \right\}$$

$$\mu(A_i) = F_X\left(\frac{i+1}{2^n}\right) - F_X\left(\frac{i}{2^n}\right)$$

The union of these disjoint sets is :

$\lim\limits_{n \to \infty} \bigcup\limits_{i=1}^{n} A_i \ni (a, b]$ which is our original interval.

Upon covering $[a, b]$ with these sets we can now sum up the measures associated with each disjoint subinterval as

$$\sum_{i=1}^{\infty} P_r\{A_i\} = \lim_{n \to \infty} \sum_{i=1}^{\infty} P_r\left\{\frac{i}{2^n} < X < \frac{i+1}{2^n}\right\}$$

This sum will be denoted as

$$\int_{\Omega} dP = \lim_{n \to \infty} \sum_{i=1}^{\infty} P_r\left\{\frac{i}{2^n} < X < \frac{i+1}{2^n}\right\}$$

$$= \lim_{n \to \infty} \sum_{i=1}^{\infty} \left[F_X\left(\frac{i+1}{2^n}\right) - F_X\left(\frac{i}{2^n}\right)\right] = 1$$

In a similar fashion, we can define the expectation operator via a Lebesgue integral

$$E\{X\} = \int_{\Omega} X \, dP = \lim_{n \to \infty} \sum_{i=1}^{\infty} \frac{i}{2^n} P_r\left\{\frac{i}{2^n} \leq X \leq \frac{i+1}{2^n}\right\}$$

$$= \lim_{n \to \infty} \sum_{i=1}^{\infty} \left(\frac{i}{2^n}\right)\left[F_X\left(\frac{i+1}{2^n}\right) - F_X\left(\frac{i}{2^n}\right)\right]$$

The difference between the Riemann Integral and the Lebesgue measure is that in the Riemann Integral the domain of the function is subdivided while the Lebesgue integral divides the range into intervals.

ECE - 541, Fall 2004

Weak Law of Large Numbers

Consider n samples drawn from a population with unknown mean μ_x and variance σ_x^2. We want to estimate μ_x via the sample mean estimate:

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} X_i ,$$

where X_i represent n i.i.d samples with parameter (μ_x, σ_x^2)

First let us look at the mean value of this estimator:

$$E\left\{\hat{\mu}_n\right\} = E\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} E\{X_i\}$$

$$= \mu_x$$

$$\Longrightarrow \quad E\left\{\hat{\mu}_n\right\} - \mu_x = 0$$

$\Longrightarrow$ $\hat{\mu}_n$ is an unbiased estimator of μ_x

Let us now look at the variance of $\hat{\mu}_n$.

$$\mathrm{Var}\{\hat{\mu}_n\} = \mathrm{Var}\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\{X_i\} = \frac{\sigma_x^2}{n}$$

$$\Rightarrow \quad E\left\{(\hat{\mu}_n - \mu_x)^2\right\} = \frac{\sigma_x^2}{n} = \sigma_{\hat{\mu}_n}^2$$

From the Chebyshev Inequality

$$P_r\left\{|\hat{\mu}_n - \mu_x| > \delta\right\} \leqslant \frac{\sigma_{\hat{\mu}_n}^2}{\delta} = \frac{\sigma_x^2}{n\delta}$$

$$\Rightarrow \quad \lim_{n\to\infty} P_r\left\{|\hat{\mu}_n - \mu_x| > \delta\right\} = \lim_{n\to\infty} \frac{\sigma_x^2}{n\delta} = 0$$

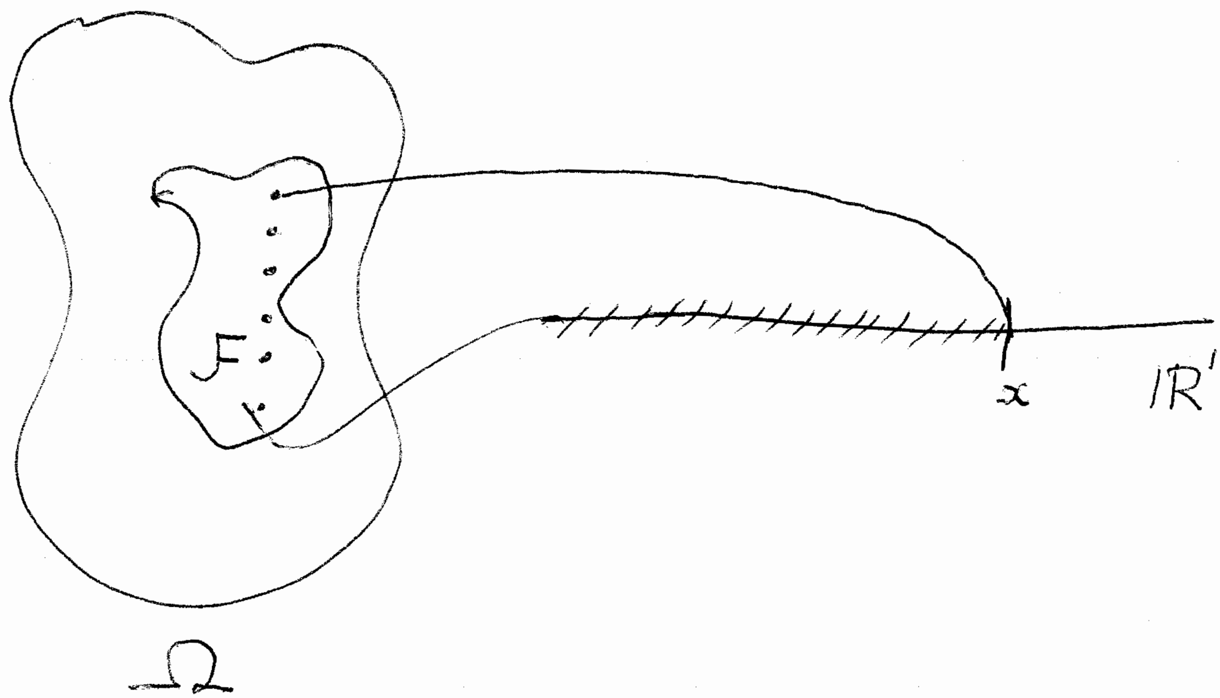
or alternatively

$$\lim_{n\to\infty} P_r\left\{|\hat{\mu}_n - \mu_x| < \delta\right\} = 1$$

$\Rightarrow$ Probability that sample mean with a infinite sample size converges to population mean is 1

$\Rightarrow$ $\hat{\mu}_n$ converges to μ_x in probability

Notion of a Random Variable



Ω: Set of outcomes

$\mathcal{F}$: Borel field or σ-algebra of measurable events

$(-\infty, x]$: Interval on real line (Borrel set on $\mathbb{R}$)

If $(\Omega, \mathcal{F})$ is a measurable space and X is a transformation defined on $\Omega \to \mathbb{R}^1$ we call X to be a $\mathcal{F}$-measurable transformation if for every $x \in \mathbb{R}^1$

$$X^{-1}((-\infty, x]) \in \mathcal{F}, \text{ i.e,}$$

$(-\infty, x]$ needs to map back to an event

For example consider a dart throwing experiment, where the transformation $X(\omega)$ is defined via :

$$X(\omega) = \begin{cases} 1, & \omega \in D \\ 0, & \omega = \text{'miss'} \end{cases}$$

where $D = \left\{ (x, y) \in \mathbb{R}^2 \ni x^2 + y^2 \leq 1 \right\}$

for $x < 0$ $\quad X^{-1}((-\infty, x])) = \phi$ since no event maps to x

for $0 \leq x < 1$ $\quad X^{-1}((-\infty, x]) = \text{'miss'}$

for $1 \leq x < \infty$ $\quad X^{-1}((-\infty, x]) = \Omega$

$$X^{-1}((-\infty, x]) \in \{ \phi, \text{'miss'}, \Omega \}$$

The minimal σ algebra here is

$$\mathcal{F}_M = \{ \phi, \text{'miss}, \Omega, D \}$$

The sigma algebra is in essence a collection of all valid logical questions regarding the experiment