Chapter 9. Integration with Respect to a Probability Measure1

Background

Let $(\Omega, \mathcal{A}, P)$ be a probability space.
We want to define the expectation, or what is equivalent, the "integral", of general $r.v.$.
We have of course already done this for $r.v.s$ on a countable space Ω.
The general case (for arbitrary Ω) is more delicate.

Definition 9.1

1. A $r.v.$ X is called **simple** if it takes on only a finite number of values and hence can be written in the form

$$X = \sum_{i=1}^{n} a_i I_{A_i} \tag{1}$$

Where $a_i \in \mathbb{R}$, and $A_i \in \mathcal{A}, 1 \leq i \leq n$

若 $A_k \in \mathcal{F}, k = 1, 2, \ldots, n$两两不交，且
$\cup_{k=1}^{n} A_k = \Omega, a_k \in \hat{\mathcal{R}}^{(1)}, k = 1, 2, \ldots, n$
则称函数

$$X(w) = \sum_{k=1}^{n} a_k I_{A_k}(w) \quad w \in \Omega$$

为 $(\Omega, \mathcal{F})$上的简单函数

- Such an X is clearly measurable;　(Why?) 思考这里为什么X是可测的

$$\begin{aligned} & \qquad\qquad\qquad \forall B \in \mathcal{B}^k \\ X^{-1}(B) &= \{w : X(w) \in B\} && \text{逆象的定义} \\ &= \cup_{a_k \in B} \{w : X(w) = a_k\} && \text{简单函数的定义} \\ &= \cup_{a_k \in B} A_k && \text{简单函数的定义} \\ &\in \mathcal{F} && \text{简单函数中} A_k \in \mathcal{F} \Rightarrow \cup_{A_k \in B} A_k \in \mathcal{F} \end{aligned}$$

由定理8.1可知X可测。

- Conversely, if X is measurable and takes on the values $a_1, \ldots, a_n$ it must have the representation (1) with $A_i = \{X = a_i\}$;
- A simple $r.v.$ has of course many different representation of the form (1).

2. If X is simple, its **expectation** (or "integral" with respect to P) is the number

$$E\{X\}=\sum_{i=1}^{n}a_iP(A_i) \tag{2}$$

- This is also written $\int X(w)P(dw)$ and even more simply $\int XdP$;
- A little algebra shows that $E\{X\}$ does not depend on the particular representation (1) chosen for X. 练习:

Exercise

Let $(\Omega,\mathcal{A},P)$ be a probability space
Let $X:\Omega\to\mathbb{R}$ be such that it admits two representations

$$X=\sum_{i=1}^{n}a_iI_{A_i}\quad\text{and}\quad X=\sum_{j=1}^{m}b_jI_{B_j}$$

where $a_i,b_j\in\mathbb{R}$, and $A_i,B_j\in\mathcal{A}$ for all i,j. Show that

$$\sum_{i=1}^{n}a_iP(A_i)=\sum_{j=1}^{m}b_jP(B_j)$$

First, prove that $\cup_{i=1}^{n}A_i=\cup_{j=1}^{m}B_j$.

Assume $\begin{matrix} a_i\neq 0 & A_i\cap A_j=\emptyset \\ b_j\neq 0 & B_i\cap B_j=\emptyset \end{matrix}(i\neq j)$

$$\begin{array}{ll}
\hline
\forall w\in\cup_{i=1}^{n}A_i & \forall w\in\cup_{j=1}^{m}B_j \\
\exists i_0\in\{1,2,\dots,n\} & \exists j_0\in\{1,2,\dots,n\} \\
s.t.\ X(w)=a_{i_0}\neq 0 & s.t.\ X(w)=b_{j_0}\neq 0 \\
so\ w\in\cup_{j=1}^{m}B_j & so\ w\in\cup_{i=1}^{n}A_i \\
\quad else\ X(w)=0 & \quad else\ X(w)=0 \\
\therefore\ \cup_{i=1}^{n}A_i\subset\cup_{j=1}^{m}B_j & \therefore\ \cup_{j=1}^{m}B_j\subset\cup_{i=1}^{n}A_i \\
\Downarrow & \\
\cup_{j=1}^{m}B_j=\cup_{i=1}^{n}A_i & \\
\hline
\end{array}$$

Second, if $A_iB_j\neq\emptyset$, for $w\in A_iB_j$ $X(w)=a_i=b_j$

$$\begin{array}{ll}
\hline
X=\sum_{i=1}^{n}a_iI_{A_i} & X=\sum_{j=1}^{m}b_jI_{B_j} \\
\quad=\sum_{i=1}^{n}a_iI_{A_i\cap(\cup_{i=1}^{n}A_i)} & \quad=\sum_{j=1}^{m}b_jI_{B_j\cap(\cup_{j=1}^{m}B_j)} \\
\quad=\sum_{i=1}^{n}a_iI_{A_i\cap(\cup_{j=1}^{n}B_j)} & \quad=\sum_{j=1}^{m}b_jI_{B_j\cap(\cup_{i=1}^{n}A_i)} \\
\quad=\sum_{i=1}^{n}a_iI_{\cup_{j=1}^{m}A_iB_j} & \quad=\sum_{j=1}^{m}b_jI_{\cup_{i=1}^{n}A_iB_j} \\
\quad=\sum_{i=1}^{n}\sum_{j=1}^{m}a_iI_{A_iB_j} & \quad=\sum_{i=1}^{n}\sum_{j=1}^{m}b_jI_{A_iB_j} \\
\Downarrow & \\
\text{for } w\in A_iB_j\neq\emptyset, X(w)=a_i=b_j & \\
\hline
\end{array}$$

如果 $A_iB_j=\emptyset$ ，$a_iP(A_iB_j)=b_jP(A_iB_j)=0$，不影响计算

Last, Prove $EX = EY$

$$
\begin{array}{ll}
\hline
EX = \sum_{i=1}^{n} a_i P(A_i) & EY = \sum_{j=1}^{m} b_j P(B_j) \\
= \sum_{i=1}^{n} a_i P(A_i \cap (\cup_{i=1}^{n} A_i)) & = \sum_{j=1}^{m} b_j P(B_j \cap (\cup_{j=1}^{m} B_j)) \\
= \sum_{i=1}^{n} a_i P(A_i \cap (\cup_{j=1}^{m} B_j)) & = \sum_{j=1}^{m} b_j P(B_j \cap (\cup_{i=1}^{n} A_i)) \\
= \sum_{i=1}^{n} a_i P(\cup_{j=1}^{m} A_i B_j) & = \sum_{j=1}^{m} b_j P(\cup_{i=1}^{n} A_i B_j) \text{ pairwise disjoint} \\
= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i P(A_i B_j) & = \sum_{i=1}^{n} \sum_{j=1}^{m} b_j P(A_i B_j) \\
\Downarrow \because a_i = b_j & \\
EX = EY & \\
\hline
\end{array}
$$

Remark [测度与概率—2.3节 期望与积分 - 知乎 (zhihu.com)]

- Let X, Y be two simple $r.v.s$ and β a real number. We clearly have

  $X = \sum_{i=1}^{n} a_i I_{A_i} \quad Y = \sum_{j=1}^{m} b_j I_{B_j}$
  $EX = \sum_{i=1}^{n} a_i P(A_i) \quad EY = \sum_{j=1}^{m} b_j P(B_j)$

  - $E\{\beta X\} = \beta E\{X\}$;

    $E\{\beta X\} = \sum_{i=1}^{n} \beta a_i P(A_i) = \beta \sum_{i=1}^{n} a_i P(A_i) = \beta E\{X\}$

  - $E\{X+Y\} = E\{X\} + E\{Y\}$;

    $$
    \begin{aligned}
    E\{X+Y\} &= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) P(A_i B_j) \\
    &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i P(A_i B_j) + \sum_{i=1}^{n} \sum_{j=1}^{m} b_j P(A_i B_j) \\
    &= \sum_{i=1}^{n} a_i P(A_i) + \sum_{j=1}^{m} b_j P(B_j) \\
    &= E\{X\} + E\{Y\}
    \end{aligned}
    $$

  - **If $X \leq Y$,then $E\{X\} \leq E\{Y\}$.**

    $$
    \begin{aligned}
    & X \leq Y \Rightarrow a_i \leq b_j \\
    & E\{X\} = \sum_{i=1}^{n} a_i P(A_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i P(A_i B_j) \\
    & \qquad \leq \sum_{i=1}^{n} \sum_{j=1}^{m} b_j P(A_i B_j) \\
    & \qquad = \sum_{j=1}^{m} b_j P(B_j) \\
    & \qquad = E\{Y\}
    \end{aligned}
    $$

- Thus, expectation is linear on the vector space of all simple $r.v.s$.

  因此，对于向量空间中的简单随机变量，期望是线性的。

- Next, we define expectation for positive $r.v.s$.
  定义正的随机变量

  For X positive,

  - By this, we assume that X may take all values in $[0,\infty]$, including $+\infty$;
    在这种假设下，随机变量X的取值为$[0,\infty]$
  - This innocuous extension is necessary for the coherence of some of our further results.
    这种无害的扩展对于我们某些进一步结果的一致性是必需的。

    Let

$$E\{X\}=\sup\{E\{Y\}: Y \text{ a simple r.v. with } 0\le Y\le X\} \tag{3}$$

  - This supremum always exists in $[0,\infty]$.
    这个上确界在 $[0,\infty]$ 上总是存在的

    Since expectation is a positive operator on the set of simple $r.v.'s$,
    既然期望是在简单随机变量集合上的正的运算

    it is clear that the definition above for $E\{X\}$ coincides with Definition 9.1.
    则上面的关于期望的定义和定义9.1是一致的

    > 定义9. 1里面的关于期望的定义为 $E\{X\}=\sum_{i=1}^n a_iP(A_i)$
    > 思考这里的一致性是为什么？

Remark

- Note that $E\{X\}\ge 0$,but we can have $E\{X\}=\infty$,even when X is never equal $+\infty$.
  注意到 $E\{X\}\ge 0$,但是我们可以有 $E\{X\}=\infty$，即使随机变量X不等于无穷。
- Finally, let X be an arbitrary $r.v.$.
  最后，令 X 为任意的随机变量

  Let $X^+=max(X,0)\quad X^-=-min(X,0)$.

  Then

$$X=X^+-X^-\quad |X|=X^++X^-$$

  and X^+,X^- are positive $r.v.s$.

Definition 9.2

- A $r.v.$ X has a **finite expectation** (is "integrable") if both $E\{X^+\}$ and $E\{X^-\}$ are finite.
  若 $E\{X^+\}$ 和 $E\{X^-\}$ 都是有限的，则随机变量 X 期望有限（也叫做可积）

  In this case, its expectation is the number 期望是两数之和

$$E\{X\}=E\{X^+\}+E\{X^-\} \tag{4}$$

also written $\int X(w)dP(w)$ or $\int XdP$

- If $X>0$ then $X^-=0$ and $X^+=X$ and, since obviously $E\{0\}=0$,this definition coincides with (3)

如果 $X>0$，则 $X^{-}=0$、$X^{+}=X$、$E\{0\}=0$，在这种情况下，和式(3)一致

We write $\mathcal{L}^1$ to denote the set of all integrable $r.v.s.$ (Sometimes we write $\mathcal{L}^1(\Omega,\mathcal{A},P)$ to remove any possible ambiguity.)
用 $\mathcal{L}^1$ 代表所有可积的随机变量的集合

- A $r.v.$ X admits an expectation if $E\{X^+\}$ and $E\{X^-\}$ are not both equal to $+\infty$.
  随机变量X有期望，则 $E\{X^+\}$ 和 $E\{X^-\}$ 不都等于正无穷。
  - Then the expectation of X is still given by (4), with the conventions $+\infty+a=+\infty$ and $-\infty+a=-\infty$ when $a\in\mathbb{R}$.
    则X的期望还是和(4)式一样，因为 $+\infty+a=+\infty$、$-\infty+a=-\infty$ $a\in\mathbb{R}$.
  - If $X\geq 0$ this definition again coincides with (3)
    如果 $X\geq 0$，则(4)和(3)是一致的
  - Note that if X admits an expectation, then $E\{X\}\in[-\infty,+\infty]$, and X is integrable if and only if its expectation is finite.
    如果X有期望，则 $E\{X\}\in[-\infty,+\infty]$,
    X 可积 $\Leftrightarrow$ 期望有限

Remark 9.1

When Ω is finite or countable we have thus two different definitions for the expectation of a $r.v.$ X, the one above and the one given in Chapter 5.
当 Ω 是有限或可数，则有两个不同的关于随机变量 X 的期望定义，一个是上面给出的，另一个是第五章给出的

In fact these two definitions coincides: it is enough to verify this for a simple $r.v.$ X, and in this case the formulas (5.1) and (9.2) are identical.
事实上，这两种定义是一致的

$$E\{X\}=\sum_{j\in T'} jP(X=j) \tag{5.1}$$

$$E\{X\}=\sum_{i=1}^{n} a_i P(A_i) \tag{9.2}$$

这个留作练习，下次课讲

Theorem 9.1

- (a) $\mathcal{L}^1$ is a vector space, and expectation is a linear map on $\mathcal{L}^1$, and it is also positive ($i.e. X\geq 0\Rightarrow E\{X\}\geq 0$).
  $\mathcal{L}^1$ 是一个向量空间，期望是$\mathcal{L}^1$上的线性映射，它也是正的。

If further $0\leq X\leq Y$ are two $r.v.s$ and $Y\in\mathcal{L}^1$ and $E\{X\}\leq E\{Y\}$.

- (b) $X\in\mathcal{L}^1$ iff $|X|\in\mathcal{L}^1$ and in this case $|E\{X\}|\leq E\{|X|\}$.

  In particular any bounded $r.v.$ is integrable.
- (c) If $X=Y$ almost surely $(s.a.)$, then $E\{X\}=E\{Y\}$.

  $$X=Y \text{ a.s. if } P(X=Y)=P(\{w:X(w)=Y(w)\})=1$$

- (d) (Monotone convergence theorem): 单调收敛定理

If the $r.v.s$ X_n are positive and increasing $a.s.$ to X, then $\lim_{n\to\infty} E\{X_n\} = E\{X\}$ (even if $E\{X\} = \infty$).

- (e) (Fatou's lemma):

  If the $r.v.s$ X_n satisfy $X_n > Y$ $a.s.$ ($Y \in \mathcal{L}^1$), all n, we have

$$E\left\{\lim_{n\to\infty} \inf E\{X_n\}\right\} \leq \lim_{n\to\infty} \inf E\{X_n\}$$

In particular, if $X_n \geq 0$ $a.s.$ then

$$E\left\{\lim_{n\to\infty} \inf X_n\right\} \leq \lim_{n\to\infty} \inf E\{X_n\}$$

- (f) (Lebesgue's dominated convergence theorem): 勒贝格控制收敛定理

  If the $r.v.s$ X_n converge $a.s.$ to X and if $|X_n| \leq Y$ $a.s. \in \mathcal{L}^1$, all n,

  then $X_n \in \mathcal{L}^1, X \in \mathcal{L}^1$ and $E\{X_n\} \to E\{X\}$.

Statement

- The $a.s.$ equality between $r.v.s$ is clearly an equivalence relation, and two equivalent (i.e. almost surely equal) r.v.s have the same expectation:
  在随机变量间的几乎必然相等时一个等价关系，两个几乎必然相等的随机变量具有相同的表示。

  Thus :

  one can define a space L^1 by considering " $\mathcal{L}^1$ modulo this equivalence relation"

- In other words, an element of L^1 is an equivalence class, that is a collection of all $r.v.$ in $\mathcal{L}^1$ which are pairwise $a.s.$ equal.
  L^1 是一个相等的类，是所有在 $\mathcal{L}^1$ 上的两两相等的类

- In view of (c) above, one may speak of the "expectation" of the equivalence class (which is the expectation of any one element belonging to this class).
  鉴于上面的（c），可以说等价类的“期望”（对属于该类的任何一个元素的期望）。

- Since further the addition of $r.v.s$ or the product of a $r.v.$ by a constant preserve $a.s.$ equality, the set L^1 is also a vector space.
  随机变量的加法或者乘法by a constant preserve $a.s.$ equality，则 L^1 集合也是一个向量空间。

  Therefore, we commit the (innocuous) abuse of identifying a $r.v.$ with its equivalence class, and commonly write $X \in L^1$ instead of $X \in \mathcal{L}^1$.

- If $1 \leq p < \infty$, we define $\mathcal{L}^p$ to be the space of $r.v.s$ such that $|X|^p \in \mathcal{L}^1$;

  L^p is defined analogously to L^1.
  L^p和L^1的定义类似

  That is, L^p is $\mathcal{L}^p$ modulo the equivalence relation "almost surely".

- Put more simply, two elements of $\mathcal{L}^p$ that are $a.s.$ equal are considered to be representatives of one element of L^p.

Two auxiliary results.

Result 1

For every positive r.v. X there exists a sequence $\{X_n\}_{n\geq 1}$ of positive simple r.v.s which increases toward X as n increases to infinity.

An example of such a sequence is given by 为什么要定义为 $\frac{k}{2^n}$

$$X_n(w) = \begin{cases} \frac{k}{2^n} & if\ \frac{k}{2^n} \le X(w) < \frac{k+1}{2^n}\ and\ 0 \le k \le n2^n - 1 \\ n & if\ X(w) \ge n \end{cases}$$

Result 2

If X is a positive r.v., and if $\{X_n\}_{n\geq 1}$ is any sequence of positive simple r.v.s increasing to X, then $E\{X_n\}$ increases to $E\{X\}$.