STAT 7200

Introduction to Advanced Probability
Lecture 7

Taylor R. Brown

"A First Look at Rigorous Probability Theory" (Jeffrey Rosenthal)
Sections 2.5 (continued), 2.6, 3.1, and 3.2

Extension Theorem

Theorem 1

The Extension Theorem *Let $\mathcal{J}$ be a semialgebra of subsets of Ω,* **P** *a function from $\mathcal{J}$ to [0,1] with the following properties:*

a) $\mathbf{P}(\emptyset) = 0, \mathbf{P}(\Omega) = 1$.

b) $\mathbf{P}(\bigcup_{i=1}^k A_i) \geq \sum_{i=1}^k \mathbf{P}(A_i)$ *whenever* $A_1, \ldots, A_k \in \mathcal{J}$, $\bigcup_{i=1}^k A_i \in \mathcal{J}$, *and* $A_1, \ldots, A_k$ *are pairwise disjoint (finite superadditivity).*

c) $\mathbf{P}(A) \leq \sum_n \mathbf{P}(A_n)$ *whenever* $A, A_1, A_2, \ldots \in \mathcal{J}$, *and* $A \subseteq \bigcup_n A_n$ *(countable monotonicity).*

Then there is a σ-algebra $\mathcal{M} \supseteq \mathcal{J}$ and a proper probability measure $\mathbf{P}^*$ on $\mathcal{M}$ so that $\mathbf{P}^*(A) = \mathbf{P}(A)$ for all $A \in \mathcal{J}$.

Variation of Extension Theorem

Proposition 2

In the original extension theorem, the finite superadditivity condition and the countable monotonicity condition of $\mathbf{P}$ *can be replaced by the following countable additivity condition:*
$\mathbf{P}(\bigcup_n A_n) = \sum_n \mathbf{P}(A_n)$ *for disjoint* $A_1, A_2, \ldots \in \mathcal{J}$ *with* $\bigcup_n A_n \in \mathcal{J}$*.*

Uniqueness of Extension Theorem

Theorem 3 (Proposition 2.5.7)

Uniqueness of Extension *In the extension theorem (or variation), the extended probability measure* $\mathbf{P}^*$ *over* $\mathcal{M}$ *is unique in the sense that: For* $\sigma-$*algebra* $\mathcal{F}$ *so that* $\mathcal{J} \subseteq \mathcal{F} \subseteq \mathcal{M}$ *and another probability measure* $\mathbf{Q}$ *over* $\mathcal{F}$ *so that* $\mathbf{Q}(A) = \mathbf{P}(A)$ *for all* $A \in \mathcal{J}$*. Then* $\mathbf{Q}(A) = \mathbf{P}^*(A)$ *for all* $A \in \mathcal{F}$*.*

- **Proof:** For any $A \in \mathcal{F}$

$$
\begin{aligned}
\mathbf{P}^*(A) &= \inf_{A_1, A_2, \ldots \in \mathcal{J}, A \subseteq \bigcup_i A_i} \sum_i \mathbf{P}(A_i) = \inf_{A_1, A_2, \ldots \in \mathcal{J}, A \subseteq \bigcup_i A_i} \sum_i \mathbf{Q}(A_i) \\
&\geq \inf_{A_1, A_2, \ldots \in \mathcal{J}, A \subseteq \bigcup_i A_i} \mathbf{Q}\left(\bigcup_i A_i\right) (\text{ countable subadditivity }) \\
&\geq \inf_{A_1, A_2 \ldots, \in \mathcal{J}, A \subseteq \bigcup_i A_i} \mathbf{Q}(A) (\text{ by monotonicity }) = \mathbf{Q}(A).
\end{aligned}
$$

Uniqueness of Extension Theorem: continued

- **Proof (continued):** The previous derivation shows that $\mathbf{P}^*(A) \geq \mathbf{Q}(A)$ for any $A \in \mathcal{F}$. Similarly, $\mathbf{P}^*(A^c) \geq \mathbf{Q}(A^c)$. But as the probability of complement equals 1 minus the probability, we have $\mathbf{P}^*(A) \leq \mathbf{Q}(A)$, thus $\mathbf{P}^*(A) = \mathbf{Q}(A)$. The extension is unique over $\mathcal{F}$.

Corollary 4 (Proposition 2.5.8)

Let $\mathcal{J}$ be a semi-algebra and $\mathcal{F}$ be the $\sigma -$ algebra generated by $\mathcal{J}$. Let $\mathbf{P}$ and $\mathbf{Q}$ be two probability measures over $\mathcal{F}$, so that $\mathbf{P}(A) = \mathbf{Q}(A)$ for any $A \in \mathcal{J}$. Then $\mathbf{P}(A) = \mathbf{Q}(A)$ for any $A \in \mathcal{F}$.

Corollary 5 (2.5.9)

Let $\mathbf{P}$ and $\mathbf{Q}$ be two probability measures over $\mathcal{B}$, the collection of Borel sets, so that $\mathbf{P}((-\infty, x]) = \mathbf{Q}((\infty, x])$ for any $x \in \mathbf{R}$. Then $\mathbf{P}(A) = \mathbf{Q}(A)$ for any $A \in \mathcal{B}$.

Infinite Number of Coin Tossing

- The sample space of tossing a fair coin infinite number of times can be denoted as: $\Omega = \{(r_1, r_2, r_3, \ldots) : r_i = 0 \text{ or } 1\}$.
- Each outcome in this sample space consists of infinite number of tosses, and each toss equals 0 or 1 with probability 0.5. Then intuitively the probability of each outcome should be 0. However, just as "the probability of $X = x$ should equal 0 if $X \sim Unif$", this result does not help us much in understanding this particular sample space.
- Denote $A_{a_1 a_2 \ldots a_n}$ ($a_i = 0$ or 1) as the event that the results of the first n tosses are exactly $a_1, a_2, \ldots, a_n$, then the collection $\mathcal{J} = \{A_{a_1 a_2 \ldots a_n} : n \in \mathbf{N}, a_i = 0 \text{ or } 1\} \bigcup \{\emptyset, \Omega\}$ is a semi-algebra. The probability function $\mathbf{P}$ over $\mathcal{J}$ can be defined as $\mathbf{P}(A_{a_1 a_2 \ldots a_n}) = 1/2^n$. And we can verify that $\mathbf{P}$ satisfies the variation of extension theorem.
- By the extension theorem, we may extend both $\mathcal{J}$ and $\mathbf{P}$ to a proper probability triple.
- This probability triple is actually equivalent to the uniform measure (Lebesgue Measure) as each $x \in [0, 1]$ can be represented as: $x = \sum_{k=1}^{\infty} \frac{a_k}{2^k}$ in a binary representation.

Product Measure

- The extension theorem is not limited to one-dimensional sample spaces. We just used it on infinite-dimensional coin-flip spaces, and we can also use it to define a uniform measure over $[0, 1] \times [0, 1]$.
- We may construct the semi-algebra as the collection of all the rectangles (may be closed or open on any of the four borders), and define $\mathbf{P}$ as the area of any rectangle. We then verify the conditions of the extension theorem as we had done for the uniform measure over [0,1] and apply the extension theorem to construct a probability triple.

Product Measure: continued

- Suppose we have two probability measures $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$. To define a probability measure over $\Omega_1 \times \Omega_2$, we may choose $\mathcal{J}$ as:

$$\mathcal{J} = \{A \times B : A \in \mathcal{F}_1, B \in \mathcal{F}_2\}$$

  and define $\mathbf{P}(A \times B) = \mathbf{P}(A) \times \mathbf{P}(B)$.
- It is quite easy to verify that $\mathcal{J}$ is a semi-algebra.
- We will show that $\mathbf{P}(A \times B)$ is countably additive later. Then we may apply the extension theorem to show that it would be possible to construct a product measure based on two marginal measures.

Random Variable: Definition

- On the probability triple $(\Omega, \mathcal{F}, \mathbf{P})$, we define a random variable $X : \Omega \to \mathbf{R}$ if
  $\{\omega : X(\omega) \leq x\} \in \mathcal{F}$ for any $x \in \mathbf{R}$ or
  $\{\omega : X^{-1}(B)\} \in \mathcal{F}$ for any $B \in \mathcal{B}$.
- The second condition certainly implies the first condition. To see why the first condition implies the second condition, define $\mathcal{A} = \{A \subseteq \mathbf{R} : X^{-1}(A) \in \mathcal{F}\}$. $\mathcal{A}$ is a σ-algebra. Moreover, for any $x \in \mathbf{R}$, $(-\infty, x] \in \mathcal{A}$. Last, by definition of the Borel σ-algebra, $\mathcal{B} \subseteq \mathcal{A}$. So, for any $B \in \mathcal{B}$, $X^{-1}(B) \in \mathcal{F}$.
- X is a "measurable" function from Ω to $\mathbf{R}$. Generally speaking, we say a function from $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ to $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ is measurable if the inverse image of any measurable set in Ω_2 is also measurable in Ω_1.

Proposition 3.1.5

- The indicator of a measurable set A (written $1_A(\omega)$) is a random variable. Take any $B \in \mathcal{B}$. 1 can be in B and/or 0 can be in B, or neither. $X^{-1}(B)$ is then either A, A^c, Ω or $\emptyset$, all of which are in $\mathcal{F}$.
- $X + c$ and cX are random variables if X is.
- X^2 is an rv if X is. $\{X^2 \leq y\} = \{-\sqrt{y} \leq X \leq \sqrt{y}\} \in \mathcal{F}$.

Proposition 3.1.5

- If X and Y are r.v.s., then $X + Y$ is still random variable as: $\{X + Y \leq z\} = \bigcup_{r \text{ is rational}} \{X \leq r\} \cap \{Y \leq z - r\}$
- If $Z_1, Z_2, \dots$ are r.v.s., and $\lim_{n \to \infty} Z_n(\omega)$ exists for every ω , then $Z = \lim_{n \to \infty} Z_n$ is also a random variable since : $\{Z \leq z\} = \bigcap_{m=1}^{\infty} \{Z \leq z + \frac{1}{m}\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{Z_n \leq z + \frac{1}{m}\}$

\- which is to say, if the limit of Z_n, Z, is smaller than z, then for any (first intersection) small number such as $1/m$, we can find a large number N (first union), so that for all (second intersection) $n \geq N$, $Z_n \leq z + \frac{1}{m}$.

Independence: Definition

- A collection of events $\{A_\alpha\}_{\alpha\in I}$ (can be finite, countable or uncountable) are independent if and only if for all finite subsets $(\alpha_1, \ldots, \alpha_J) \subseteq I$, $P(\bigcap_{j=1}^{J} A_{\alpha_j}) = \prod_{j=1}^{J} P(A_{\alpha_j})$.

- It is possible to have any two of the three events to be independent but the three events together are not. For instance, throw a fair dice twice, let $A = \{\text{First toss is even}\}$, $B = \{\text{Second toss is odd}\}$ and $C = \{\text{Sum of the first and second dices is even}\}$.

- A collection of random variable $\{X_\alpha\}_{\alpha\in I}$ are independent if and only if for all finite subsets $(\alpha_1, \ldots, \alpha_J) \subseteq I$, and all Borel sets $S_1, \ldots, S_J$, $P(X_{\alpha_1} \in S_1, \ldots, X_{\alpha_J} \in S_J) = P(X_{\alpha_1} \in S_1) \ldots P(X_{\alpha_J} \in S_J)$.
- If random variable X and Y are independent, then $f(X)$ and $g(Y)$ are also independent for measurable functions f and g.