Probability Theory

TR 1:00 – 2:30 CB 120

Grading Scheme

- 40% Assignments
- 20% Midterm Exam (February 25)
- 40% Final Exam (Same as comprehensive exam)

Material Covered:

- Probability Spaces
- Random Variables and Distributions
- Expected Value
- Convergence of Random Variables
- Limit Theorems
- Characteristic Functions
- Conditional Probability and Expectation
- Martingales

Probability Spaces

Motivation The probability of an event

Examples Probability that

(α) 2 heads in 3 coin tosses

(β) 3 coin tosses land HHT

(γ) At least 3 cm of snow on January 31.

(δ) Two random real numbers $x, y \in [0,10]$ such that $x+y > 4$.

Definition: A probability measure $P: \mathcal{F} \to [0,1]$, where $\mathcal{F}$ denotes the set of events, assigns a value in $[0,1]$ to an event.

To represent an event, we start with a sample space Ω representing the set of all possible events that can happen, in the context of interest. An event is a subset of Ω

Examples:

(α) $\Omega_1 = \{0,1,2,3\}$, where n represents the number of heads in 3 tosses.

(β) $\Omega_2 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

using Ω_2, event in (α) corresponds to $\{HHT, HTH, THH\}$. Using Ω_1, event in (β) corresponds to $\{2\}$.

(γ) $\Omega_3 = [0, \infty)$, $x \in \Omega_3$ represents exactly x cm of snow on January 31

Some Rules For $P: \mathcal{F} \to [0,1]$

(i) $P(\Omega) = 1$ and $P(\emptyset) = 0$

(ii) $P(A^c) = 1 - P(A)$. More generally, if $A_1 \cap A_2 = \emptyset$ then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.

Even more generally, if $A_1, \ldots, A_n$ are pairwise disjoint,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \quad (\text{finite additivity})$$

We want countable additivity. That is, if $(A_n)_{n=1}^{\infty}$ is a collection of disjoint events,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Examples: Consider Ω_2 from above. For $A \subset \Omega_2$,

$$P(A) = \sum_{x \in A} \{x\}$$

e.g. $P(\{2 \text{ heads}\}) = P(\{HHT\}) + P(\{HTH\}) + P(\{THH\})$

If each sequence is equally likely as for a fair coin then we get the above probability.

Note that countable (finite or infinite) additivity holds whenever Ω is countable.

In (γ) and (δ) from above, Ω is uncountable.

Example: Consider $\Omega_4 = [0, 10]$ with a point chosen uniformly at random. Here, this means for $a < b \in [0, 10]$

$$P([a,b]) = \frac{b-a}{10}$$

Also, $P(\{x\}) = \frac{x-x}{10} = 0$.

Theorem: There is no way to define a countably additive function $P : \mathcal{F} \to [0,1]$, where $\mathcal{F}$ denotes the subsets of $[0,10]$, such that $\forall a<b \in [0,10]$

$$P([a,b]) = \frac{b-a}{10}$$

Proof: (in notes)

The idea to avoid this to reduce the domain of P to a collection $\mathcal{S}$ of subsets of $\Omega = [0,10]$

<u>Need</u>: (i) every subinterval is in $[0,10]$

(ii) $\Omega, \emptyset \in \mathcal{S}$

(iii) $A_1, A_2 \in \mathcal{S} \Rightarrow A_1 \cup A_2 \in \mathcal{S}$, $A_1 \setminus A_2 \in \mathcal{S}$

(iv) $(A_n)_{n=1}^{\infty} \in \mathcal{S} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$ and $\bigcap_{n=1}^{\infty} A_n \in \mathcal{S}$

A collection of sets satisfying (ii) and (iii) is called an algebra, and those satisfying (ii)–(iv) are called a σ-algebra.

Undergraduate Probability Distributions

(α) Poisson, Binomial, ... (Discrete)

Ω is countable, $f(x) = P(\{x\}) \geq 0$ s.t.

$\sum_{x \in \Omega} P(\{x\}) = 1$. For $A \subset \Omega$, $P(A) = \sum_{x \in A} P(\{x\})$

We can take $\mathcal{S}$ to be $P(\Omega)$, the collection of all subsets of Ω.

(β) Continuous, Normal, exponential, ...

$\Omega = \mathbb{R}$ or $[0, \infty)$, ...

$\exists$ probability density function $g(x) \geq 0$

$P(\mathbb{R}) = \int_{-\infty}^{\infty} g(x)\,dx = 1$

$P([a,b]) = \int_a^b g(x)\,dx$

$P(\{a\}) = 0$.

For $A \subset \mathbb{R}$, $P(A) = \int_A g(x)\,dx$