01/23/06 Fundamentals of Probability Theory

Homework 1 posted due February 6 (Monday)

Probability space: \( \Omega \)

Discrete: coin flips
terrorist status (yes/no) in some population

Can you see me now?

Continuous: location of particle \((\mathbb{R}^3)\)

Abstract: Financial markets over the next year
Status of New Orleans after an Atlantic hurricane

Elements \( \omega \) describe complete information for one possible outcome

An "event" \( A \) is a set of possible outcomes \( A \subseteq \Omega \)
\[ \text{Prob}(A) = \leq \text{Prob}(\omega) \quad \text{for discrete state spaces} \]

The \text{Prob}(\omega) must be specified by model.

"Prob" behaves like a positive measure,

\[ \text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B) \]

\[ = \text{Prob}(A \text{ or } B) \quad \text{if the events } A \text{ and } B \text{ are mutually exclusive} \]

\[ (A \land B = \emptyset) \]

\[ \text{Prob}(A) \geq 0 \]

\[ \text{Prob}(\emptyset) = 0 \]

\[ \text{Prob}(\Omega) = 1 \]

\[ \text{Prob}(\Omega \setminus A) = \text{Prob}(\Omega) - \text{Prob}(A) \]

\[ = 1 - \text{Prob}(A) \]

\[ \text{Prob}(\text{not } A) \]

Technically, the probability measure "Prob" may only be defined for some collection of subsets of \( \Omega \).

- Important for continuous and abstract prob spaces.
In dependence: Events $A$ and $B$ have no

Two events $A$ and $B$ are independent if the outcome of $A$ happening or not happening has no connection to whether $B$ happens.

$$\text{Prob}(A \text{ and } B) = \text{Prob}(A) \cdot \text{Prob}(B)$$

$$\text{Prob}(A \cap B)$$

More general notion of connection of two events

Conditional probability:

$$\text{Prob}(A \mid B) = \frac{\text{Prob}(A \text{ and } B)}{\text{Prob}(B)}$$

$$\text{Prob}(A \text{ happens given that } B \text{ happens}) \text{ (true in general)}$$
If uniform probability
\[ \text{Prob}(A) = \frac{1}{4} \]
\[ \text{Prob}(B) = \frac{1}{4} \]
\[ \text{Prob}(A \text{ and } B) = \frac{1}{8} \]
\[ \text{Prob}(\overline{A} \mid B) = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2} \]

Statistical independence
\[ \implies \text{Prob}(A \mid B) = \text{Prob}(A) \]
if \(A, B\) independent.

**Probability**

Random variables are "measurable" functions on probability space.

The range of a random variable is called the *state space*.

Measurable means, for random variables \(X\), with discrete state space, that \(\text{Prob}(X = x)\) is well defined for any value \(x \in S\) (state space).
More formally:

\[ X : \mathcal{N} \to S \]

\[ \mathrm{Prob}(X = x) = \bigvee_{\omega \in \mathcal{N}} \mathrm{Prob}(\omega) \quad \text{where} \quad X(\omega) = x \]

To be a little more precise, we make the probability measures on the various spaces more precise:

For \( A \subseteq \mathcal{N} \)

\[ \mathrm{Prob}(A) \equiv P(A) \]

For \( B \subseteq \mathcal{S} \)

\[ B \subseteq S \]

\[ \mathrm{Prob}(B) = P_X(B) \]

\[ P_X(B) = \bigvee_{\omega \in \mathcal{N}} \mathrm{Prob}(\omega) \quad \text{where} \quad X(\omega) \in B \]

\[ P_X \]

is a measure induced on \( \mathcal{S} \) by pushing forward the measure \( P \) on \( \mathcal{N} \) through the function \( X : \mathcal{N} \to \mathcal{S} \).
\( P \) is a measure defined on some "measurable" collection of subsets of \( \mathbb{R} \).

For discrete state spaces, a practical way to work is to define:

\[
P_x = P(X = x) = \text{Prob}(X = x)
\]

\[
P_X(B) = \sum_{x \in B} P_x
\]

This allows one to work with the set of positive reals \( \{ p_x \}_{x \in S} \) rather than measures.

Practical into about random variable \( X \) (discrete state space still)

Probability distribution: \( \{ p_x \}_{x \in S} \)

- Complete statistical info about \( X \)
Useful condensations of info about $X$:

Mean (average or expected value):

$$
\mu_X = \langle X \rangle = \mathbb{E} X
$$

$$
= \sum_{x \in S} x \cdot p_x
$$

More generally for any function $f$ defined on $S$

$$
\langle f(X) \rangle = \sum_{x \in S} f(x) \cdot p_x
$$

Average/expected value of $f(X)$

Variance:

$$
\sigma_X^2 = \langle (X - \mu_X)^2 \rangle
$$

reflects how much $X$ tends to deviate from mean value.
Simple rules for manipulating $< >$: behave like sums + integrals

For example, can formally view:

This is a linear operation:

$$< a f(X) + b g(X) >$$

$$= a < f(X) > + b < g(X) >$$

for $a, b \in \mathbb{F}$ (deterministic)

Parameters which are not random behave like parameters independent of integration variable.

- can factor out of integral.

$$\int y x^2 \, dx = y \int x^2 \, dx$$

Example calculation:

$$\mu_x^2 = < (X - \mu_X)^2 >$$

$$= < X^2 - 2 X \mu_X + \mu_X^2 >$$

$$= < X^2 > - 2 \mu_X < X > + \mu_X^2 < 1 >$$

because $\mu_X$ is not random.
\[ \begin{align*}
\sigma^2 &= \langle \bar{X}^2 \rangle - \mu^2 \\
\mu^2 &= \langle \bar{X} \rangle - \mu
\end{align*} \]

Standard deviation: \( \sigma_{\bar{X}} = \sqrt{\sigma^2} \)

- Typical departure from mean

Further information about random variables from higher moments:

\[ \langle X^n \rangle = \sum_{x \in S} x^n p_x \quad \text{for } n \geq 1, 3, \ldots \]

Actually, "cumulants" are a better way to organize information on modified moments.
Third moment (cumulant): skewness
Fourth moment/cumulant: kurtosis
- rapidness of decay in tails

Slower decay
Larger kurtosis
Examples of random variables with discrete state space:

1) Binomial distribution

\[ S = \{0, 1, 2, \ldots, N\} \]

\[ \text{Prob} (\bar{X} = j) = \binom{N}{j} p^j (1-p)^{N-j} \]

where \( p \) is a prescribed value \( 0 < p < 1 \).

\[ \binom{N}{j} = \frac{N!}{(N-j)! j!} \]

Interpretation: Bernoulli trials 
(coin flips)

- \( N \) independent trials
- \( p \) = probability of positive outcome in any given trial,

\[ \bar{X} = \# \text{ positive outcomes in } N \text{ trials} \]
2) Uniform distribution

\[ S = \{0, 1, \ldots, N\} \]

\[ \text{Prob} (X = j) = \frac{1}{N+1} \quad \text{for } 0 \leq j \leq N. \]

Equal likelihood for each outcome,

Office hours

Wed 3-4 PM
Fri 4-5 PM