Review Probability Theory

- "Basic" Probability Theory
- Laws of Large Numbers

"Basic" Probability Theory

Probability Model

- S: sample space (all possible outcomes)
  - E: event, subset of S
- $\mathcal{F}$: set of events
- P: rule for assigning probabilities to events

Examples of sample spaces:

- $S = \{w_1, ..., w_n\}$: finite number of sample points
- $S = \{w_1, w_2, ...\}$: countably infinite number of sample points
- $S = [w_l, w_u]$: uncountably infinite number of sample points

Structure of set $\mathcal{F}$

- If $A, B \in \mathcal{F}$, then
  - $A^C \in \mathcal{F}$
  - $A \cup B \in \mathcal{F}$
  - $A \cap B \in \mathcal{F}$
- If $A_1, A_2, A_3, ... \in \mathcal{F}$, then
  - $\cup_i A_i \in \mathcal{F}$
  - $\cap_i A_i \in \mathcal{F}$

Axioms for Probability Measure $P : \mathcal{F} \mapsto [0, 1]$

1. $0 \leq P(E) \leq 1, \qquad E \in \mathcal{F}$
2. $P(S) = 1$
3. If $E_1, E_2, ...$ are such that $E_j \cap E_j = \emptyset$, $i \neq j$, then

$$P(\cup_i E_i) = \sum_i P(E_i)$$

Examples:
Uniform Distribution on $S = [0, 1]$

- $P(\{w\}) = \qquad\qquad , w \in S$
- $P(E) = \qquad\qquad , E = [x, y] \subset S$

From Axioms:

- $P(A^C) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Event of Probability 1

- $E \in \mathcal{F}$
- $E \neq S$
- $P(E) = 1$

Example

Conditional Probability

- $A, B \in \mathcal{F}$, $P(B) > 0$
- $P(A \mid B) = \frac{P(AB)}{P(B)}$

Independence

- $P(AB) = P(A)P(B)$
- If A and B are independent, then

$$P(A \mid B) = P(A)$$

Conditional Independence

- $P(AB \mid C) = P(A \mid C)P(B \mid C)$
- Example: A, B are not independent, but A, B are conditionally independent.

Random Variable:

- $X : S \mapsto R = (-\infty, +\infty)$ (!!!)
- $P(X \leq x) = P(\{w \in S : X(w) \leq x\})$
  - function of real variable x
  - non-decreasing in x
  - distribution function $F_X(x)$

Examples

Probability Density

- $f_X(x) = \frac{d}{dx} F_X(x)$
- If density exists and is finite everywhere, then X is a continuous RV

Probability Mass Function for Discrete RV

- $P_X(x_i), \quad i = 1, 2, ...$

Multiple Random Variables:

- $X_1, ..., X_n$ are n RV.

Example: n dice rolls

Distribution Functions

- $F_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n)$: joint distribution function
- $F_{X_i}(x_i) = F_{X_1,X_2,...,X_n}(\infty, \infty, ..., x_i, ..., \infty)$: marginal distribution

The RV's $X_1, ..., X_n$ are independent if **for all** $x_1, ..., x_n \in \Re^N$ we have

$$F_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n) = \Pi_{i=1}^N F_{X_i}(x_i)$$

Expectations:

- Cont. RV: $E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$
- Discrete RV: $E[X] = \sum_x x P_X(x)$
- General Case: $E[X] = \bar{X} = \int_{-\infty}^{\infty} x dF_X(x)$

Note: A RV is finite with probability 1, but this does not imply that the expectation is finite.

Non-negative RV

- $X : S \mapsto [0, \infty)$
- $E[X] = \int_0^{\infty} \Big[1 - F_X(x)\Big] dx$
- $1 - F_X(x)$: complimentary distribution function

Functions of RV

- $Y = g(X)$
- $g : (-\infty, \infty) \mapsto (-\infty, \infty)$
- $Y(w) = g\big(X(w)\big), \qquad w \in S$
- $E[Y] = \int_{-\infty}^{\infty} y dF_Y(y) = \int_{-\infty}^{\infty} g(x) dF_X(x)$

Moments

- $E[X^n]$: nth moment of X
- $E[(X - \bar{X})^n]$: nth central moment X
- $VAR(X) = \sigma_X^2 = E[(X - \bar{X})^2] = E[X^2] - \bar{X}^2$

Sums of RV

- $Z = X + Y$
- If X, Y are RV's then Z is a RV
- If X, Y are independent RV's, then

$$F_Z(z) = \int_{-\infty}^{\infty} F_X(z - y) dF_Y(y)$$

- We always have

$$E[X_1 + X_2 + ... + X_n] = E[X_1] + ... + E[X_n]$$

- If $X_1, ..., X_n$ are independent RV's, then

$$E[X_1 X_2 ... X_n] = E[X_1]E[X_2]...E[X_n]$$

$$\sigma_{S_n}^2 = \sigma_{X_1}^2 + ... + \sigma_{X_n}^2, \qquad S_n = X_1 + ... + X_n$$

Laws of Large Numbers

$X_1, X_2, ..., X_n$: IID RV with finite mean $\bar{X}$

$S_n = X_1 + X_2 + ... + X_n$

Weak Law of Large Numbers

$$\lim_{n\to\infty} P\left[\left|\frac{S_n}{n} - \bar{X}\right| \geq \epsilon\right] = 0, \qquad \epsilon > 0$$

Strong Law of Large Numbers

$$\lim_{n\to\infty} P\left[\sup_{m\geq n}\left|\frac{S_m}{m} - \bar{X}\right| \geq \epsilon\right] = 0, \qquad \epsilon > 0$$

These results are very important for later

Basic Inequalities

For **Non-negative** RV:

- $P(Y \geq y) \leq \frac{E[Y]}{y}$, Markov Inequality
- If $E[Y] < \infty$, then

$$\lim_{y\to\infty} yP(Y \geq y) = 0.$$

For RV with finite mean and finite variance

- $P\big[|Z - E[Z]| \geq \epsilon\big] \leq \frac{\sigma_Z^2}{\epsilon^2}$, Chebyshev Inequality

Weak Law of Large Numbers for IID RV with finite mean and finite variance

Central Limit Theorem

$X_1, X_2, ..., X_n$: IID RV with finite mean $\bar{X}$ and finite variance σ_z^2

$S_n = X_1 + X_2 + ... + X_n$

Central Limit Theorem

$$\lim_{n\to\infty} P\left(\frac{S_n - n\bar{X}}{\sqrt(n)\sigma_x} \leq y\right) = \int_{-\infty}^{y} \frac{1}{\sqrt(2\pi)} exp\left(\frac{-x^2}{2}\right) dx.$$

$Y_n = \frac{S_n - n\bar{X}}{\sqrt(n)\sigma_x}$

- $E[Y_n] = 0$
- $\sigma_{Y_n}^2 = 1$
- $\lim_{n\to\infty} F_{Y_n}(y) = \int_{-\infty}^{y} \frac{1}{\sqrt(2\pi)} exp\left(\frac{-x^2}{2}\right) dx.$

Relative Frequency and Indicator Functions

Consider $(S, \mathcal{F}, P)$ and an event $A \in \mathcal{F}$

- Indicator Function I_A (RV)

$$I_A = \begin{cases} 1 & w \in A \\ 0 & \text{otherwise} \end{cases}$$

- $E[I_A] =$
- $\sigma_A^2 = P(A)[1 - P(A)]$
- I_{A_i}, $\quad i = 1, ..., n$, n IID RV
- Relative Frequency of A

$$\frac{\sum_{i=1}^{n} I_{A_i}}{n}$$

From weak law of large numbers

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \bar{X}\right| \geq \epsilon\right) = 0, \qquad \epsilon > 0.$$

Consistent within probability model: relative frequency of an event can be related to the probability of an event

Laws of Large Numbers

$X_1, X_2, ..., X_n$: IID RV with finite mean $\bar{X}$ and possible infinite variance

$S_n = X_1 + X_2 + ... + X_n$

Weak Law of Large Numbers

$$\lim_{n\to\infty} P\left[\left|\frac{S_n}{n} - \bar{X}\right| \geq \epsilon\right] = 0, \qquad \epsilon > 0$$

Strong Law of Large Numbers

$$\lim_{n\to\infty} P\left[\sup_{m\geq n}\left|\frac{S_m}{m} - \bar{X}\right| \geq \epsilon\right] = 0, \qquad \epsilon > 0$$