Review of Probability

Models ch. 1-2
Simulation ch. 2 ← nice, concise review

Outline

1 - Sample spaces & events
- Conditional probability
- Bayes' Formula
- Independence

So basically all of 461/661 in 1+ lecture!

2 - Random Variables
  - Discrete / Continuous RVs
  - Common Distributions (Table on website)
  - Expectation, Variance
  - Joint Distributions
  - Moment generating functions
  - LLN, CLT, other limit theorems

Sample Space & Events

Consider an experiment (outcome not known in advance).

def: Let S be the sample space of the experiment: the set of all possible outcomes.

Examples:

- Toss a coin: $S = \{H, T\}$
- Toss a coin twice: $S = \{HH, HT, TH, TT\}$
  ordered, here all outcomes are equally likely (all outcomes have prob. $\frac{1}{4}$)
- Roll a die:
  $S = \{1, 2, 3, 4, 5, 6\}$
- Roll a pair of dice: $S = \{(i,j) : i,j \in \{1, \ldots, 6\}\}$
  $|S| = 36$ equally likely outcomes

def: Event = subset of S

e.g. event $A = \{2, 4, 6\}$ roll an even #

Operations on events (sets)

union $\cup$
intersection $\cap$
complement
mutually exclusive events (aka disjoint): $A \cap B = \phi$

Axioms of Probability

For each event A of S, the probability of A denoted $P(A)$ satisfies the following axioms:

1) $0 \leq P(A) \leq 1$

2) $P(S) = 1$

3) For any sequence of mutually exclusive events $A_1, A_2, \ldots$ (i.e. $A_i \cap A_j = \phi$ for $i \neq j$)

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \quad \text{for } n=1,2,\ldots$$

Note: $1 = P(S) = P(\underbrace{A \cup A^c}_{\text{always mut. exclusive}}) = P(A) + P(A^c)$

$\Rightarrow \boxed{P(A^c) = 1 - P(A)}$ – very useful property!

So $P(\phi) = 1 - P(S) = 0$

Note: Events A & B on S.

$A \subset B \Rightarrow \boxed{P(A) \leq P(B)}$

Addition Rule: $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$

Generalizes to Inclusion/Excl. Principle

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

— Fundamental Concept in Prob.

<u>Conditional Probability</u>

Events A & B on sample space S.

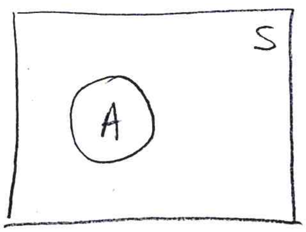
$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}} \quad \text{for } P(B) > 0$$

$$\Rightarrow \quad P(A \cap B) = P(A \mid B)\,P(B) = P(B \mid A)\,P(A)$$

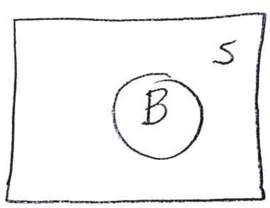
"Multiplication Rule"

<u>Intuition</u>:

universe (all possible outcomes) $= S$

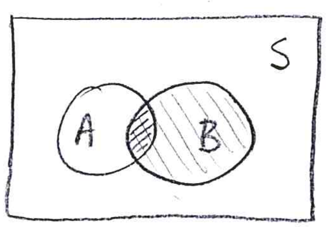


$$P(A) = \frac{\#\text{ of outcomes in } A}{\#\text{ outcomes in } S}$$



same idea for $P(B)$

Now suppose we know event B has occurred.
what's the probability A will occur? $P(A \mid B)$



Now the universe is B.
Looking for $P(A \cap B)$ in that univ. so we divide by $P(B)$

$$\therefore \; P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Example: Roll 2 dice. Observe that the 1st one is 5. What is the probability that the sum of the two dice is 7?

B : 1st die is 5 Find $P(A \mid B)$.

A : sum is 7

Possible Outcomes : $(5,1), \underbrace{(5,2)}, (5,3), (5,4), (5,5), (5,6)$

only one that sums to 7

So $P(A \mid B) = \frac{1}{6}$

OR use formula

$$\left[P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{6}{36} = \frac{1}{6} \right]$$

Q. What if A : sum is ≤ 7 ?

$P(A \cap B) = \frac{2}{36}$

$\Rightarrow P(A \mid B) = \frac{1}{3}$

Independence

def: Two events A & B are independent if

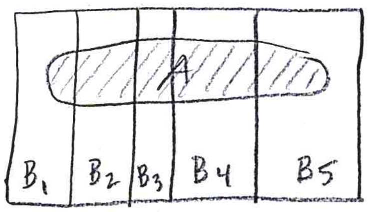
$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

Also, $\boxed{P(A \mid B) = P(A)}$.

conditioning on B has no effect on prob. of A

Bayes' Formula & Law of total prob.

def: Partitioned sample space : Sequence of events $B_1, \dots, B_n$



s.t. $B_i \cap B_j = \phi$ for $i \neq j$

and $\bigcup_{i=1}^{n} B_i = S$

Law of total probability

$$P(A) = \underbrace{\sum_{i=1}^{n} P(A \cap B_i)}_{= P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right)} = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

since $(A \cap B_i) \cap (A \cap B_j) = \phi$ for $i \neq j$ ~ mut. excl. events

Bayes' Formula

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)\,P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$ from LTP above

Inverse cond. prob

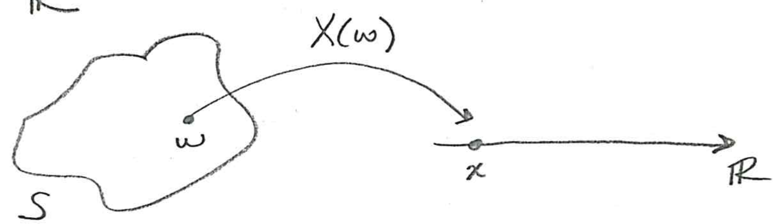
Random Variables

Recall: Probability experiment (e.g. flip a coin 10 times)
large sample space S (describe S)

Often we are interested in the value of some numerical quantity determined by the outcome of exp.
(e.g. # of heads in 10 tosses)
described by a random variable

def: A random variable (RV) X is a function from S to $\mathbb{R}$, i.e. real-valued function.

$X : S \to \mathbb{R}$



e.g. $X \in \{0, 1, 2, \dots, 10\}$ for exp. above

Discrete vs. Continuous RVs

finite or countably ∞ S

uncountable S (i.e. interval of $\mathbb{R}$)

Probability density (or (mass)) function (PDF) p_X or f_Y

or just density fct

discrete RV X: $p_X(k) = P(X=k)$ for all $k \in S$

continuous RV Y: $f_Y(y)$ s.t. $\underset{= P(Y \in [a,b])}{P(a \le Y \le b)} = \int_a^b f_Y(y)\,dy$

Cumulative distribution function (CDF) F_X or F_Y

or just distribution

discrete X: $F_X(k) = P(X \le k) \left(= \sum_{\text{all } j \le k} P(X=j) \right)$

continuous Y: $F_Y(y) = P(Y \le y) = \int_{-\infty}^{y} f_Y(t)\,dt$ for any $y \in \mathbb{R}$

Properties

PDFs
- $f(t) \geq 0 \quad \forall t$
- $\int_{-\infty}^{\infty} f(t)\,dt = 1$

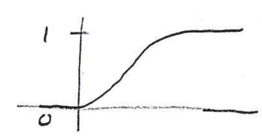
} same holds for mass function

* * probabilities are
- b/t 0 & 1
- never negative!

CDFs
- $P(a < Y < b) = F_Y(b) - F_Y(a)$
- $P(Y > a) = 1 - P(Y \le a) = 1 - F_Y(a)$

$\lim_{y \to \infty} F_Y(y) = 1$

$\lim_{y \to -\infty} F_Y(y) = 0$



Q. Relationship b/t density & dist'n functions?

density is derivative of dist'n

Common Distributions

Discrete: Bernoulli, Binomial, Geometric, Poisson, [Multinomial, Neg. Binomial]

Continuous: Uniform, Exponential, Gamma, Normal, Beta, Pareto

stopped

Bernoulli RV

Binary outcome < success / failure   e.g. coin flip

Let $X = \begin{cases} 1 & \text{if success} \\ 0 & \text{if failure} \end{cases}$

Probability mass function of X:

$$\begin{cases} P(X=1) = p & \text{- success prob} \\ P(X=0) = 1-p \end{cases}$$

$$\Rightarrow \boxed{P(X=k) = p^k (1-p)^{1-k}, \quad \text{for } k = 0,1}$$

$X \sim$ Bernoulli(p)

preview for later

$$E[X] = p$$
$$Var(X) = p(1-p)$$

Binomial RV

aka Bernoulli trials

(e.g. coin flips)

Suppose n independent trials, each of which results in success w/prob. p & failure with prob. $1-p$.

Let X = # of successes in the n trials. (e.g. # of heads in n coin flips)

$X \sim$ Binomial (n, p) $\quad$ ($X \sim \text{Bin}(n,p)$)

PMF of X: $$\boxed{P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k=0,1,\ldots,n}$$

* Also, Binomial RV is sum of n i.i.d. Bernoulli RVs

$X = \sum_{i=1}^{n} X_i$

$X \sim \text{Bin}(n,p)$

$X_i \sim \text{Bern}(p)$

Multinomial RV

Generalizes Binomial to r outcome types (like rolling an r-sided die n times)

See formula sheet for PMF

Geometric & Negative Binomial

Geometric: # trials until 1st success

Negative Binomial: # trials until r^{th} success → generalizes geometric distn

(Also, Neg. Bin $= \sum_{i=1}^{r}$ iid Geom(p))

Talk about Poisson dist'n next week!

Poisson: can be used to approx. a binomial RV when n large & p small ($\lambda = np$)