Dean Zeller
Graph Mining – Class Notes
Thursday, February 1st, 2007

Probability

Definition 1: Probability Space (Ω, F, P)

Ω – sample space: all possible results.

Example: roll a six sided die, results is $\Omega = \{1\ 2\ 3\ 4\ 5\ 6\}$

F – power set of Ω: set of all subsets.

$F = 2^{\Omega}$

$|F| = 2^{|\Omega|}$

Example: $\{\varnothing\ \{1\}\ \{2\}\ \{3\}\ \ldots\ \{6\}\ \{12\}\ \{13\}\ \ldots\ \{26\}\ \ldots\ \{66\}\ \{111\}\ \{112\}\ \ldots\ \{123456\}\}$

P – probability function: maps a subset of the sample space to a probability

$F \to R$

Definition 2: P: $F \to R$

E is an event, $E \subseteq F$

A) $0 \le P(E) \le 1$

B) $P(\Omega) = 1$

C) $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ iff $E_1 \cap E_2 = \varnothing$

$P(E_1 \cup E_2 \cup E_3 \ldots \cup E_n) = P(E_1) + P(E_2) + P(E_3) + \ldots + P(E_n)$ iff $E_i \cap E_j = \varnothing\ \forall\ i \neq j$

Lemma 1: $P((E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

Proof:

Note that $E_1 = E_1 - (E_1 \cap E_2) + (E_1 \cap E_2)$

Thus…

$P(E_1) = P(E_1 - (E_1 \cap E_2)) + P(E_1 \cap E_2)$ and

$P(E_2) = P(E_2 - (E_1 \cap E_2)) + P(E_1 \cap E_2)$

$P(E_1 \cup E_2) = P(E_1 - (E_1 \cap E_2)) + P(E_2 - (E_1 \cap E_2)) + P(E_1 \cap E_2))$

and

$P(E_1 \cup E_2) - P(E_1) - P(E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

Definition 3: Conditional Probability

Events $A, B \in F$

$P(A \mid B)$ – Probability of event A given event B is true

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

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = \underbrace{\underbrace{P(A_1) \cdot P(A_2 \mid A_1)}_{P(A_1 \cap A_2)} \cdot P(A_3 \mid A_1 \cap A_2)}_{P(A_1 \cap A_2 \cap A_3)} \cdot \ldots \cdot P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1})$$

Definition 4: Random Variables

Choose an event from the sample space Ω. A random variable X is an "observation."

X: $\Omega \to R$

Discrete, either a finite number of possibilities or countable infinite.

$X = a$ is $\{y \in \Omega \mid X(y) = a\}$ "a set in Ω in which the random value is a"

Example: roll a six-sided die twice, sum the results

Ω = {1-1, 1-2, 1-3, 1-4, 1-5, 1-6, 2-1, 2-2, 2-3, 2-4, 2-5, 2-6, 3-1, 3-2, 3-3, 3-4, 3-5, 3-6, 4-1, 4-2, 4-3, 4-4, 4-5, 4-6, 5-1, 5-2, 5-3, 5-4, 5-5, 5-6, 6-1, 6-2, 6-3, 6-4, 6-5, 6-6}

$|\Omega| = 36$

Each event has $\frac{1}{36}$ probability.

X=4 is {13, 22, 31}

$P(X=4) = \frac{3}{36} = \frac{1}{12}$

$P(X=2) = \frac{1}{36}$ {1-1}

$P(X=7) = \frac{6}{36} = \frac{1}{6}$ {1-6, 2-5, 3-4, 4-3, 5-2, 6-1}

Definition 5: Independent events

X and Y are independent iff $\forall\, x,y\ P((X=x)\cap(Y=y)) = P(X=x) \cdot P(Y=y)$

Definition 6: Expectation

The expectation of a random variable X is

$$E(X) = \sum_{i \in R(X)} i \cdot P(X = i)$$

Example: Pick 10 cards. What is the expected number of red cards? $(10 \times .5 = 5)$

Lemma 2: Linearity of Expectation

Given events $X_1, X_2, \ldots, X_k$ (mutually exclusive)

$E(X_1 + X_2 + \ldots + X_k) = E(X_1) + E(X_2) + \ldots + E(X_k)$

Example:

X: card is a spade = $\frac{13}{52}$ = .25

Y: card is a king = $\frac{4}{52}$ = .077

$P(X \cap Y)$ = P(King of spaces) = $\frac{1}{52} = .019 = \frac{13}{52} \times \frac{4}{52}$

Thus, X and Y are mutually exclusive.

Counterexample:

X: card is a face card = $\frac{12}{52}$ = .23

Y: card is a king = $\frac{4}{52}$ = .077

$P(X \cap Y) \neq \frac{1}{52}$

Thus, X and Y are not mutually exclusive.

Prove E(X + Y) = E(X) + E(Y)

|X| = i |Y| = j

$$\begin{aligned}
E(X+Y) &= \sum_{i+j \in R(X+Y)} (i+j) P((X+Y) = i+j) \\
&= \sum_{i+j \in R(X+Y)} (i+j) P(X = i \wedge Y = j) \\
&= \sum_{i+j \in R(X+Y)} \left[i \cdot P(X = i \wedge Y = j) + j \cdot P(X = i \wedge Y = j) \right] \\
&= \sum_{i \in R(X)} \sum_{j \in R(Y)} i \cdot P(X = i \wedge Y = j) + \sum_{i \in R(X)} \sum_{j \in R(Y)} j \cdot P(X = i \wedge Y = j) \\
&= E(X) + E(Y)
\end{aligned}$$

Definition 7: Variance of a random variable X

$\mathrm{Var}(X) = E(X - E(X))^2$ $E(X)$ is the mean of the data

$= E(X^2 - 2X\,E(X) + (E(X))^2)$
$= E(X^2) - E(2X\,E(X)) + E(E(X)^2)$
$= E(X^2) - 2E(X)E(X) + (E(X)^2)$

Thus $\mathrm{Var}(X) = E(X - E(X))^2 = E(X^2) - (E(X))^2$

Standard deviation: $\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Definition 8: Covariance of two random variables X, Y

Note the similarity between Covariance and Variance. Variance is the covariance of a variable with itself.

$\mathrm{Cov}(X,Y) = E((X-E(X))\cdot(Y-E(Y))$
$= E(X\cdot Y - X\cdot E(Y) - Y\cdot E(X) + E(X)\cdot E(Y))$
$= E(X\cdot Y) - E(X\cdot E(Y)) - E(Y\cdot E(X) + E(X)\cdot E(Y))$
$= E(X\cdot Y) - E(X)\cdot E(Y) + E(X)\cdot E(Y) + E(X)\cdot E(Y)$
$= E(X\cdot Y) - E(X)\cdot E(Y)$

Lemma 3: $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$

Proof:

$\mathrm{Var}(X+Y) = E(X+Y - E(X+Y))^2$
$= E((X+Y)^2 - 2(X+Y)E(X+Y) + (E((X+Y))^2)$
$= E((X+Y)^2) - (E(X+Y))^2$
$= E(X^2) + 2E(X\cdot Y) + E(Y^2) - (E(X+Y))^2$
$= E(X^2) + 2E(X\cdot Y) + E(Y^2) - (E(X) + E(Y))^2$
$= E(X^2) + 2E(X\cdot Y) + E(Y^2) - (E(X)^2 + 2E(X)E(Y) + E(Y)^2$
$= (E(X^2) - (E(X)))^2 + (E(Y^2) - (E(Y))^{2)} + (2E(X\cdot Y) - 2E(X)E(Y))$
$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$

Lemma 4: Events X,Y are independent $\rightarrow E(X\cdot Y) = E(X)\cdot E(Y)$

$\mathrm{Cov}(X,Y) = 0$

$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Example: Bernoulli distribution

$$Y = \begin{cases} 1: \text{tails} & p \\ 2: \text{heads} & 1-p \end{cases}$$

$P(Y=1) = p$
$P(Y=0) = 1 - p$

$E(Y) = p$
$\mathrm{Var}(Y) = E(Y^2) - (E(Y))^2$

Toss a coin n times: $Y_1, Y_2, \ldots Y_n$
X is the number of 1's.
$X = Y_1 + Y_2 + \ldots Y_n$
$E(X) = E(Y_1 + Y_2 + \ldots Y_n)$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$E(X) = \sum_{k=0}^{n} k\cdot P(x = k)$$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Var}(Y_1 + Y_2 + \ldots\ Y_n) = np \\ &= \mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + \ldots\ \mathrm{Var}(Y_n) \\ &= np(1-p)\end{aligned}$$