Probability

Chang-Gun Lee (cglee@snu.ac.kr)

Assistant Professor

The School of Computer Science and Engineering

Seoul National University

Sample Space

- Sample space Ω: collection of all possible experimental outcomes
  - E.g., If we roll a die, the possible outcomes are {1, 2, ..., 6}
  - E.g., If we look at (sample) a queue, the possible numbers of customers in the queue are {0,1,2,…}
  - E.g., If we sense (sample) a room temperature, the possible outcomes are [-50, +50]
- Two parts of a sample space
  - It contains a list of all the outcomes of some experiments
  - It quantifies the likelihood of each of these outcomes
- A more careful definition of a sample space
  - a set of outcomes: S
  - a function that assigns a numerical score (probability) to each outcome such that the sum of the probabilities of all the outcomes to be exactly 1: P
- Definition 29.2 (Sample Space): A sample space is a pair (*S*, *P*) where *S* is a finite, nonempty set and *P* is a function *P*: $S \rightarrow R$ such that $P(s) \geq 0$ for all $s \in S$ and

$$\sum_{s \in S} P(s) = 1$$

Sample Point

- Sample point ω: one outcome of a sample
  - E.g., In the experiment of rolling a die, “1” is a sample point
  - E.g., The first sample point of the queue = 0
  - E.g., The first sample point of the temperature = -50 degree
- Example 29.4 (Pair of dice) Two dice are tosses. Define the sample space. How many sample points the sample space has? What is the probability of a sample point (1,6)?
- Example 29.6 (Coin tossing) A fair coin is tossed five times in a row, and the sequence of HEADS and TAILS is recorded. Define the sample space. How many sample points? What is the probability of each sample point?

Events

- Event A: set of sample points
  - E.g., In the die-tossing example, we can define an event that the die will show an even number, that is Event A = {2, 4, 6}
  - E.g., Event A: queue is not empty = {1, 2, 3, …}
  - E.g., Event B: the temperature is higher than 10 degree = [+10, +50]
- Definition 30.1 (Event) Let (S,P) be a sample space. An event A is a subset of S (i.e., $A \subseteq S$). The probability of an event A, denoted $P(A)$, is

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

- Example 30.3 (Coin tossing) Let (S,P) be the sample space that models tossing a coin five times. What is the probability that we see exactly one HEAD?
  - Define the event
  - Calculate the probability of the event
- Example 30.4 (Ten dice) Ten dice are tossed. What is the probability that none of the dice shows the number 1?

Combining Events

- Events are subsets of a probability space. We can use the usual operations of set theory (e.g., union and intersection) to combine events
  - E.g., In the die-tossing example, suppose A is the event that a die shows an even number and B is the event that it shows a prime number. Then $A \cup B = \{2,4,6\} \cup \{2,3,5\} = \{2,3,4,5,6\}$
- Complement of an event, that is, the event that A does not occur

$$\overline{A} = S - A$$

Properties of Events

- Proposition 30.7: Let A and B be events in a sample space (S, P). Then

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

- Proof:???
- Proposition 30.8: Let (*S*,*P*) be a sample space and let *A* and *B* be events. We have the following:

$$\text{If } A\cap B=\phi, P(A\cup B)=P(A)+P(B)$$

$$P(A\cup B)\leq P(A)+P(B)$$

$$P(S)=1$$

$$P(\phi)=0$$

$$P(\overline{A})=1-P(A)$$

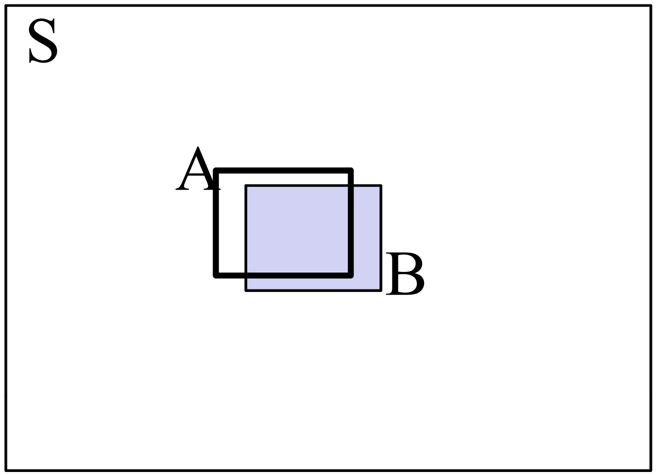
Birthday Problem

- Four people are chosen at random. What is the probability that two (or more) of them have the same birthday
  - ignore the possibility that a person might born on Feb. 29
  - it is equally likely that a person is born on any given day of the year
  - → Solution = 1.64%

- 23 people are chosen at random. What is the probability that some of them have the same birthday?
  - → Solution=50.73%

Conditional Probability (1)

- Example: Let A represent the event that a student misses the school bus. Let B represent the event that student's alarm clock malfunctions.
  - Both these events have low probability; P(A) and P(B) are small numbers
  - "What is the probability of the student missing the school bus given the fact that the alarm clock malfunctioned?" → Now it is likely the student will miss the bus!
  - We denote this probability as P(A|B): This is the probability that event A occurs given that event B occurs.



Conditional Probability (2)

- Definition 31.1 (Conditional probability) Let A and B be events in a sample space (S,P) and suppose $P(B) \neq 0$. The conditional probability $P(A|B)$, the probability of A given B, is

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

- Example 31.3: A coin is flipped five times. What is the probability that the first flip is a TAIL given that exactly three HEADS are flipped?

Independence

- Example: A coin is flipped five times. What is the probability that the first flip comes up HEADS given that the last flip comes up HEADS?
  - Let A be the event that the first flip comes up HEADS
  - Let B be the event that the last flip comes up HEADS

$$P(A) = \frac{2^4}{2^5} = 1/2, P(B) = \frac{2^4}{2^5} = 1/2, P(A \cap B) = \frac{2^3}{2^5} = 1/4$$

$$P(A \mid B) = \frac{1/4}{1/2} = 1/2$$

- Event A has nothing to do with event B, A and B are independent
- Proposition 31.4: Let A, B be events in a sample space (S,P) and suppose P(A) and P(B) are both nonzero. Then the following statements are equivalent:

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

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

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

Independent events

- Definition 31.5 (Independent events) Let A and B be events in a sample space. We say that these events are independent provided

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

- Example: A bag contains twenty balls; ten of the balls are painted red and ten are painted blue. Two balls are drawn from the bag.
  - Let A be the event that the first ball drawn is red
  - Let B be the event that the second ball is red.
- Are these two events independent?

replacement

$$P(A) = \frac{200}{400}$$

$$P(B) = \frac{200}{400}$$

$$P(A \cap B) = \frac{100}{400}$$

no-replacement

$$P(A) = \frac{10 \times 19}{20 \times 19} = \frac{1}{2}$$

$$P(B) = \frac{10 \times 10 + 10 \times 9}{20 \times 19} = \frac{1}{2}$$

$$P(A \cap B) = \frac{10 \times 9}{20 \times 19} = \frac{9}{38}$$

Independent repeated trials

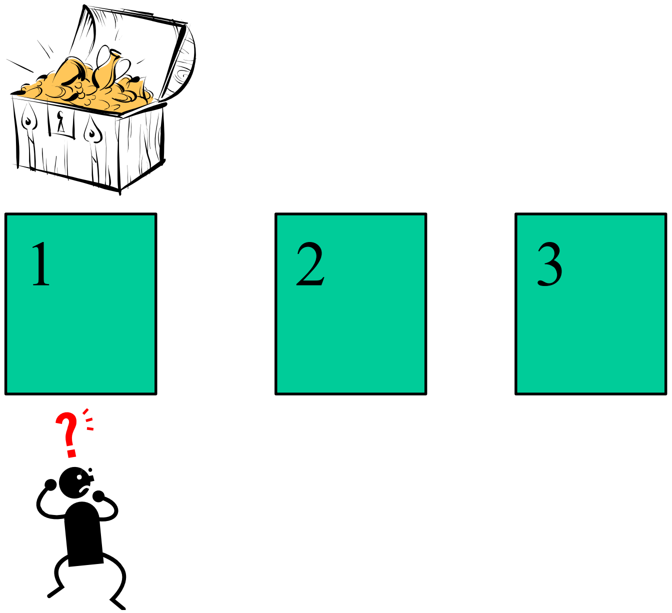
- Definition 31.6 (Repeated trials) Let (*S*,*P*) be a sample space and let *n* be a positive integer. Let S^n denote the set of all length-*n* lists of elements in *S*. Then (S^n, *P*) is the *n*-fold repeated-trial sample space in which

$$P[(s_1, s_2, \cdots, s_n)] = P(s_1)P(s_2)\cdots P(s_n)$$

- Example 31.8 Consider a sample space representing five flips of a fair coin.
  - S={HEADS,TAILS} and P(s)=1/2 for both s in S.
  - Define the sample space for the "toss-five-times" experiment
- Example 31.9 Imagine a coin that is not fairly balanced; that is, it does not turn up HEADS and TAILS with the same probabilities.
  - S={HEADS,TAILS} and P(HEADS)=p and P(TAILS)=1-p
  - If we toss this coin five times, what is the probability that we see: HHTTH?

Monty Hall Problem

- Let’s make a Deal show hosted by Monty Hall



Random Variables

- We might not be interersted in the specific outcomes in a sample space, but might be interested in some quantity derived from the outcome.
  - sum of the numbers on two dice
  - number of HEADS observed in ten throws of a fair coin
- Definition 32.1 (Random variable) A random variable is a function defined on a probability space; that is, if (S, P) is a sample space, then a random variable is a function X:S→V (for some set V)
  - E.g., X is the modular 10 of customer count in the queue (discrete)
  - E.g., Y is the room temperature in Fahrenheit (continuous)
- Example 32.2 (Pair of dice) Let (S,P) be the pair-of-dice sample space. Let X:S→N be the random variable that gives the sum of the numbers on the two dice. For example, X[(1,2)]=3, X[(5,5)]=10, and X[(6,2)]=8
- Example 32.3 (Heads minus tails) Let (S,P) be the sample space representing ten tosses of a fair coin. Let X:S→Z be the random variable that gives the number of HEADS minus the number of TAILS. For example, X(HHTHTTTTHT)= -2. We can also define random variables X_H and X_T as the number of HEADS and the number of TAILS in an outcome. For example, X_H(HHTHTTTTHT)=4 and X_T(HHTHTTTTHT)=6. Notice that $X=X_H - X_T$

Representing Events with Random Variables

- Example: if we roll a pair of dice, what is the probability that the sum of the numbers is 8?
  - Define A be the event that the two dice sum to 8; A={(2,6),(3,5),(4,4),(5,3),(6,2)}. What is P(A)?
  - Define a random variable X to be the sum of the numbers on the dice. What is the probability that X=8? We read "X=8" as an event
  - "X=8" means $\{s \in S{:}X(s)=8\}$
  - What does $P(X \geq 8)$ mean?

$$P(X \geq 8) = P(\{s \in S : X(s) \geq 8\}) = \frac{5+4+3+2+1}{36} = \frac{15}{36}$$

- Ten flips of a fair coin. X_H is the number of HEADS and X_T is the number of TAILS. What is the probability that there are at least four HEADS and at least four TAILS?

$$P(X_H \geq 4 \wedge X_T \geq 4) = P(4 \leq X_H \leq 6) = \frac{\binom{10}{4} + \binom{10}{5} + \binom{10}{6}}{2^{10}}$$

Binomial random variable

- Unfair coin. Suppose this coin produces HEADS with probability *p* and TAILS with probability 1-*p*. The coin is flipped n times. Let *X* denote the number of times that we see HEADS?

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

- Think of expansion of $(p+q)^n$

Independent Random Variables

- Recall the pair-of-dice sample space.
- $X_1(s)$ is the number on the first die
- $X_2(s)$ is the number on the second die
- $X = X_1+X_2$

- Knowledge of X_2 tells us some information about X.
- However, knowledge of X_2 tells us nothing about X_1.
- The events "$X_1=a$" and "$X_2=b$" are independent for all *a* and *b*.

- Definition 32.6 (Independent random variables) Let (*S*,*P*) be a sample space and let *X* and *Y* be random variables defined on (*S*,*P*). We say that *X* and *Y* are independent if, for all *a*, *b*,

$$P(X = a \text{ and } Y = b) = P(X = a)P(Y = b)$$

Expectation

- Definition 33.1 (Expectation) Let *X* be a real-valued random variable defined on a sample space (*S*,*P*). The expectation (or the expected value or mean value) of *X* is

$$E(X) = \sum_{s \in S} X(s)P(s)$$

- Example: Suppose we roll a pair of dice. Let X be the sum of the numbers on the two dice. What is the expected value of X?
  - 36 sample points
  - 36 additions
- Proposition 33.4 Let (S,P) be a sample space and let X be a real-valued random variable defined on S. Then

$$E(X) = \sum_{a \in R} aP(X = a)$$

- Proof: ???
- Apply Proposition 33.4 to the above example
- Example 33.6: A random variable X is defined as the absolute value of the difference of the numbers on the two dice. What is the expected value of X?

Linearity of Expectation (1)

- Proposition 33.7: Suppose X and Y are real-valued random variables defined on a sample space (S,P). Then

$$E(X+Y)=E(X)+E(Y)$$

- Proof:???
- What is the expected value of the sum of the numbers on the two dice?
  - $Z = X_1+X_2$
  - $E(Z) = E(X_1)+E(X_2)=3.5+3.5=7$
- More complicated example: A basket holds 100 chips that are labeled with the numbers 1 through 100. Two chips are drawn at random from the basket (without replacement). What is the expected value of their sum, X?
  - by original definition: summation involves 9900 terms
  - by Proposition 33.4: X can vary from 3 to 199, summation involves 197 terms
  - by Proposition 33.7: Let X_1 be the number on the first chip and X_2 the number on the second chip.

$$E(X_1)=E(X_2)=\frac{1+2+\cdots+100}{100}=\frac{5050}{100}=50.5$$

$$E(X)=E(X_1+X_2)=101$$

- Proposition 33.7 holds even for dependent random variables!

Linearity of Expectation (2)

- Proposition 33.9: Let *X* be a real-valued random variable on a sample space (*S*,*P*) and let *c* be a real number. Then

$$E(cX) = cE(X)$$

- Theorem 33.10 (Linearity of expectation) Suppose *X* and *Y* are real-valued random variables on a sample space (*S*,*P*) and suppose *a* and *b* are real numbers. Then

$$E(aX + bY) = aE(X) + bE(Y)$$

- Example: A coin is tossed 10 times. Let X be the number of times we observe TAILS immediately after seeing HEADS. What is the expected value of X?
  - Let X_1 be the random variable whose value is one if the first two tosses are HEADS-TAILS and is zero othewise
  - $X_2, \ldots, X_9$ are similarly defined.
  - $X = X_1+X_2+\ldots+X_9$
  - $E(X_k) = 0\ P(X_k=0) + 1\ P(X_k=1)$, $P(X_k=1)=1/4$
  - $E(X)=9/4$

Product of Random Variables (1)

- Question: E(XY)=E(X)E(Y)?
- Example 33.13: A fair coin is tossed twice. Let X_H be the number of HEADS and let X_T be the number of TAILS observed. Let $Z=X_HX_T$. What is E(Z)?
  - $E(X_H)=E(X_T)=1$, so $E(X_HX_T)=1$?
  - E(Z)=0P(Z=0)+1(Z=1)=0*2/4 + 1*2/4=1/2
  - $E(X_HX_T)\neq E(X_H)E(X_T)$
- Theorem 33.14 Let X and Y be independent, real-valued random variables defined on a sample space (S,P). Then

$$E(XY)=E(X)E(Y)$$

- Proof: ???

$$E(Z)=\sum_{a\in R} aP(Z=a)$$

$$=\sum_{a\in R} a\left[\sum_{b,c\in R:bc=a} P(X=b\wedge Y=c)\right]=\sum_{a\in R} a\left[\sum_{b,c\in R:bc=a} P(X=b)P(Y=c)\right]$$

$$=\sum_{a\in R}\left[\sum_{b,c\in R:bc=a} aP(X=b)P(Y=c)\right]=\sum_{a\in R}\left[\sum_{b,c\in R:bc=a} bcP(X=b)P(Y=c)\right]$$

$$=\sum_{b,c\in R:bc} bcP(X=b)P(Y=c)=\sum_{b\in R}\left[\sum_{c\in R} bP(X=b)cP(Y=c)\right]$$

$$=\sum_{b\in R} bP(X=b)\left[\sum_{c\in R} cP(Y=c)\right]=\left[\sum_{b\in R} bP(X=b)\right]\left[\sum_{c\in R} cP(Y=c)\right]$$

$$E(X)E(Y)$$

Product of Random Variables?

- Question: If X and Y satisfy E(XY)=E(X)E(Y), then may we conclude that X and Y are independent?
  - NO

Variance

- Consider the following three random variables

$$X = \begin{cases} -2 & \text{with probability } 1/2 \\ 2 & \text{with probability } 1/2 \end{cases}$$

$$Y = \begin{cases} -10 & \text{with probability } 0.001 \\ 0 & \text{with probability } 0.998 \\ 10 & \text{with probability } 0.001 \end{cases}$$

$$Z = \begin{cases} -5 & \text{with probability } 1/3 \\ 0 & \text{with probability } 1/3 \\ 5 & \text{with probability } 1/3 \end{cases}$$

- How do we measure the level of "spread out" of an random variable?
- Definition 33.16 (Variance) Let X be a real-valued random variable on a sample space (S,P). The *variance* of X is

$$\text{Var}(X) = E\left[(X - E(X))^2\right]$$

- Proposition 33.19: Let X be a real-valued random variable. Then

$$\text{Var}(X) = E\left[X^2\right] - E\left[X\right]^2$$

- Proof?

Variance of Binomial random variable

- An unfair coin is flipped n times. The coin produces HEADS with probability p and TAILS with probability 1-p. Let X denote the number of times we see HEADS. We have E(X)=np. What is the variance of X?

- Solution
  - Let X_j=1 if the j-th flip comes up HEADS and X_j=0 if the j-th flip comes up TAILS
  - $X=X_1+X_2+\ldots+X_n$

$$X^2 = [X_1 + X_2 + \cdots + X_n]^2$$

$$= \sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j$$

$$E[X^2] = E\left[\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j\right]$$

$$= \sum_{i=1}^{n} E[X_i^2] + \sum_{i \neq j} E[X_i X_j]$$

$$= \sum_{i=1}^{n} E[X_i] + \sum_{i \neq j} E[X_i] E[X_i]$$

$$= np + n(n-1)p^2$$

$$\mathrm{Var}[X^2] = E[X^2] - E[X]^2$$

$$= np + n(n-1)p^2 - (np)^2$$

$$= np + n^2p^2 - np^2 - n^2p^2$$

$$= np(1-p)$$

Homework

- 29.1, 29.2, 29.5
- 30.2, 30.7, 30.14
- 31.1, 31.13, 31.23
- 32.1, 32.3, 32.7
- 33.2, 33.4, 33.6, 33.10, 33.15, 33.16