CMSC 474, Introduction to Game Theory

Introduction to Probability Theory*

Mohammad T. Hajiaghayi
University of Maryland

*: Some slides are adopted from slides by Rong Jin

Outline

- Basics of probability theory
- Bayes' rule
- Random variable and distributions: Expectation and Variance

Definition of Probability

- ***Experiment***: toss a coin twice
- ***Sample space***: possible outcomes of an experiment
  - S = {HH, HT, TH, TT}
- ***Event***: a subset of possible outcomes
  - A={HH}, B={HT, TH}
- ***Probability of an event*** : an number assigned to an event Pr(A)
  - Axiom 1: 0<= Pr(A) <= 1
  - Axiom 2: Pr(S) = 1, Pr(Ø)= 0
  - Axiom 3: For two events A and B, Pr(A∪B)= Pr(A)+Pr(B)-Pr(A∩B)
  - Proposition 1: Pr(~A)= 1- Pr(A)
  - Proposition 2: For every sequence of disjoint events $\Pr(\bigcup_i A_i) = \sum_i \Pr(A_i)$

Joint Probability

- For events A and B, **joint probability** Pr(AB) (also shown as Pr(A ∩ B)) stands for the probability that both events happen.
- Example: A={HH}, B={HT, TH}, what is the joint probability Pr(AB)?

  Zero

Independence

- Two events ***A and B are independent*** in case

$$\Pr(AB) = \Pr(A)\Pr(B)$$

- A set of events $\{A_i\}$ is ***independent*** in case

$$\Pr\left(\bigcap_i A_i\right) = \prod_i \Pr(A_i)$$

Independence

- Two events ***A and B are independent*** in case

$$\Pr(AB) = \Pr(A)\Pr(B)$$

- A set of events $\{A_i\}$ is ***independent*** in case

$$\Pr\left(\bigcap_i A_i\right) = \prod_i \Pr(A_i)$$

- Example: Drug test

	Women	Men
Success	200	1800
Failure	1800	200

A = {A patient is a Woman}

B = {Drug fails}

Will event A be independent from event B ?

Pr(A)=0.5, Pr(B)=0.5, Pr(AB)=9/20

Independence

- Consider the experiment of tossing a coin twice
- Example I:
  - A = {HT, HH}, B = {HT}
  - Will event A independent from event B?
- Example II:
  - A = {HT}, B = {TH}
  - Will event A independent from event B?
- Disjoint ≠ Independence
- If A is independent from B, B is independent from C, will A be independent from C?

Not necessarily, say A=C

Conditioning

- If A and B are events with Pr(A) > 0, the ***conditional probability of B given A*** is

$$\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}$$

Conditioning

- If A and B are events with Pr(A) > 0, the ***conditional probability of B given A*** is

$$\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}$$

- Example: Drug test

	Women	Men
Success	200	1800
Failure	1800	200

A = {Patient is a Woman}

B = {Drug fails}

Pr(B|A) = ?

Pr(A|B) = ?

Conditioning

- If A and B are events with Pr(A) > 0, the ***conditional probability of B given A*** is

$$\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}$$

- Example: Drug test

	Women	Men
Success	200	1800
Failure	1800	200

A = {Patient is a Woman}

B = {Drug fails}

Pr(B|A) = 18/20

Pr(A|B) = 18/20

- Given A is independent from B, what is the relationship between Pr(A|B) and Pr(A)?

Pr(A|B)= P(A)

Conditional Independence

- Event A and B are ***conditionally independent given C*** in case

  $$\Pr(AB|C)=\Pr(A|C)\Pr(B|C)$$

- A set of events $\{A_i\}$ is conditionally independent given C in case

  $$\Pr(\cap_i A_i|C) = \Pi_i \Pr(A_i|C)$$

Conditional Independence (cont'd)

- Example: There are three events: A, B, C
  - Pr(A) = Pr(B) = Pr(C) = 1/5
  - Pr(A,C) = Pr(B,C) = 1/25, Pr(A,B) = 1/10
  - Pr(A,B,C) = 1/125
  - Whether A, B are independent? $1/5*1/5 \neq 1/10$
  - Whether A, B are conditionally independent given C?

    $Pr(A|C)= (1/25)/(1/5)=1/5$, $Pr(B|C)= (1/25)/(1/5)=1/5$

    $Pr(AB|C)=(1/125)/(1/5)= 1/25=Pr(A|C)Pr(B|C)$
- A and B are independent ≠ A and B are conditionally independent

Outline

- Basics of probability theory
- Bayes’ rule
- Random variables and distributions: Expectation and Variance

Bayes' Rule

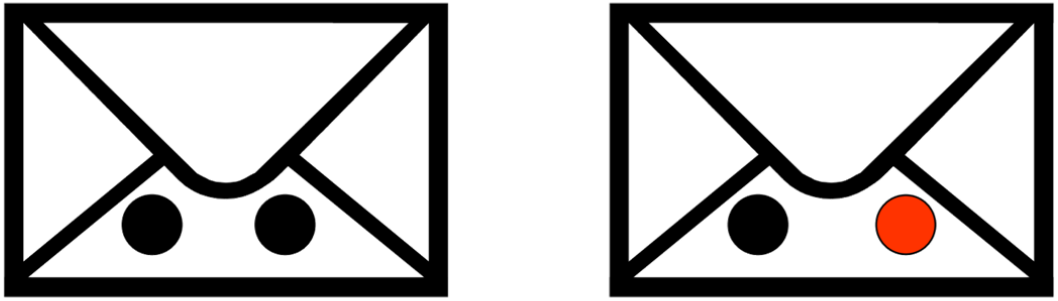
- Given two events A and B and suppose that Pr(A) > 0. Then

$$\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)} = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)}$$

- Why do we make things this complicated?
  - Often P(B|A), P(A), P(B) are easier to get
  - Some names:
    - **Prior P(A)**: probability before any evidence
    - **Likelihood P(B|A)**: assuming A, how likely is the evidence
    - **Posterior P(A|B)**: conditional prob. after knowing evidence
    - **Inference**: deriving unknown probability from known ones

Inference with Bayes' Rule: Example

- In a bag there are two envelopes
  - one has a red ball (worth $100) and a black ball
  - one has two black balls. Black balls worth nothing



- You randomly grabbed an envelope, randomly took out one ball – it's black.
- At this point you're given the option to switch the envelope. To switch or not to switch?

Inference with Bayes' Rule: Example

- E: envelope, 1=(R,B), 2=(B,B)
- B: the event of drawing a black ball
- P(E|B) = P(B|E)*P(E) / P(B)
- We want to compare P(E=1|B) vs. P(E=2|B)
- P(B|E=1) = 0.5, P(B|E=2) = 1
- P(E=1)=P(E=2)=0.5
- P(B)=3/4 (it in fact doesn't matter for the comparison)
- P(E=1|B)=1/3, P(E=2|B)=2/3
- After seeing a black ball, the posterior probability of this envelope being 1 (thus worth $100) is smaller than it being 2
- Thus you should switch

Bayes' Rule: More Complicated

- Suppose that B_1, B_2, … B_k form a partition of S:

$$B_i \bigcap B_j = \varnothing; \quad \bigcup_i B_i = S$$

Suppose that $\Pr(B_i) > 0$ and $\Pr(A) > 0$. Then

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i)\Pr(B_i)}{\Pr(A)}$$

Bayes' Rule: More Complicated

- Suppose that $B_1, B_2, \ldots B_k$ form a partition of S:

$$B_i \bigcap B_j = \varnothing; \quad \bigcup_i B_i = S$$

Suppose that $\Pr(B_i) > 0$ and $\Pr(A) > 0$. Then

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i)\Pr(B_i)}{\Pr(A)}$$

$$= \frac{\Pr(A \mid B_i)\Pr(B_i)}{\sum_{j=1}^{k} \Pr(AB_j)}$$

Bayes' Rule: More Complicated

- Suppose that B_1, B_2, … B_k form a partition of S:

$$B_i \bigcap B_j = \varnothing; \quad \bigcup_i B_i = S$$

Suppose that $Pr(B_i) > 0$ and $Pr(A) > 0$. Then

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i)\Pr(B_i)}{\Pr(A)}$$

$$= \frac{\Pr(A \mid B_i)\Pr(B_i)}{\sum_{j=1}^{k} \Pr(AB_j)}$$

$$= \frac{\Pr(A \mid B_i)\Pr(B_i)}{\sum_{j=1}^{k} \Pr(B_j)\Pr(A \mid B_j)}$$

Outline

- Basics of probability theory
- Bayes' rule
- Random variable and probability distribution: Expectation and Variance

Random Variable and Distribution

- A ***random variable X*** is a numerical outcome of a random experiment
- The ***distribution*** of a random variable is the collection of possible outcomes along with their probabilities:
  - Discrete case: $\Pr(X = x) = p_\theta(x)$
  - Continuous case: $\Pr(a \le X \le b) = \int_a^b p_\theta(x)dx$
- The ***support*** of a discrete distribution is the set of all *x* for which $\Pr(X=x) > 0$
- The ***joint distribution*** of two random variables *X* and *Y* is the collection of possible outcomes along with the joint probability $\Pr(X=x, Y=y)$.

Random Variable: Example

- Let S be the set of all sequences of three rolls of a die. Let X be the sum of the number of dots on the three rolls.
- What are the possible values for X?
- Pr(X = 3) = 1/6*1/6*1/6=1/216,
- Pr(X = 5) = ?

Expectation

- A random variable X~Pr(X=x). Then, its expectation is

$$E[X] = \sum_x x \Pr(X = x)$$

  - In an empirical sample, x_1, x_2,…, x_N,

$$E[X] = \frac{1}{N}\sum_{i=1}^{N} x_i$$

- Continuous case: $E[X] = \int_{-\infty}^{\infty} x p_\theta(x) dx$
- In the discrete case, expectation is indeed the average of numbers in the support weighted by their probabilities
- Expectation of sum of random variables

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

Expectation: Example

- Let S be the set of all sequence of three rolls of a die. Let X be the sum of the number of dots on the three rolls.
- Exercise: What is E(X)?

- Let S be the set of all sequence of three rolls of a die. Let X be the product of the number of dots on the three rolls.
- Exercise: What is E(X)?

Variance

- The variance of a random variable X is the expectation of $(X-E[X])^2$:

$$Var(X)=E[(X-E[X])^2]$$

$$=E[X^2+E[X]^2-2XE[X]]=$$

$$=E[X^2]+E[X]^2-2E[X]E[X]$$

$$=E[X^2]-E[X]^2$$

Bernoulli Distribution

- The outcome of an experiment can either be success (i.e., 1) and failure (i.e., 0).
- Pr(X=1) = p, Pr(X=0) = 1-p, or

$$p_\theta(x) = p^x (1-p)^{1-x}$$

- E[X] = p, Var(X) = $E[X^2]-E[X]^2 = p-p^2$

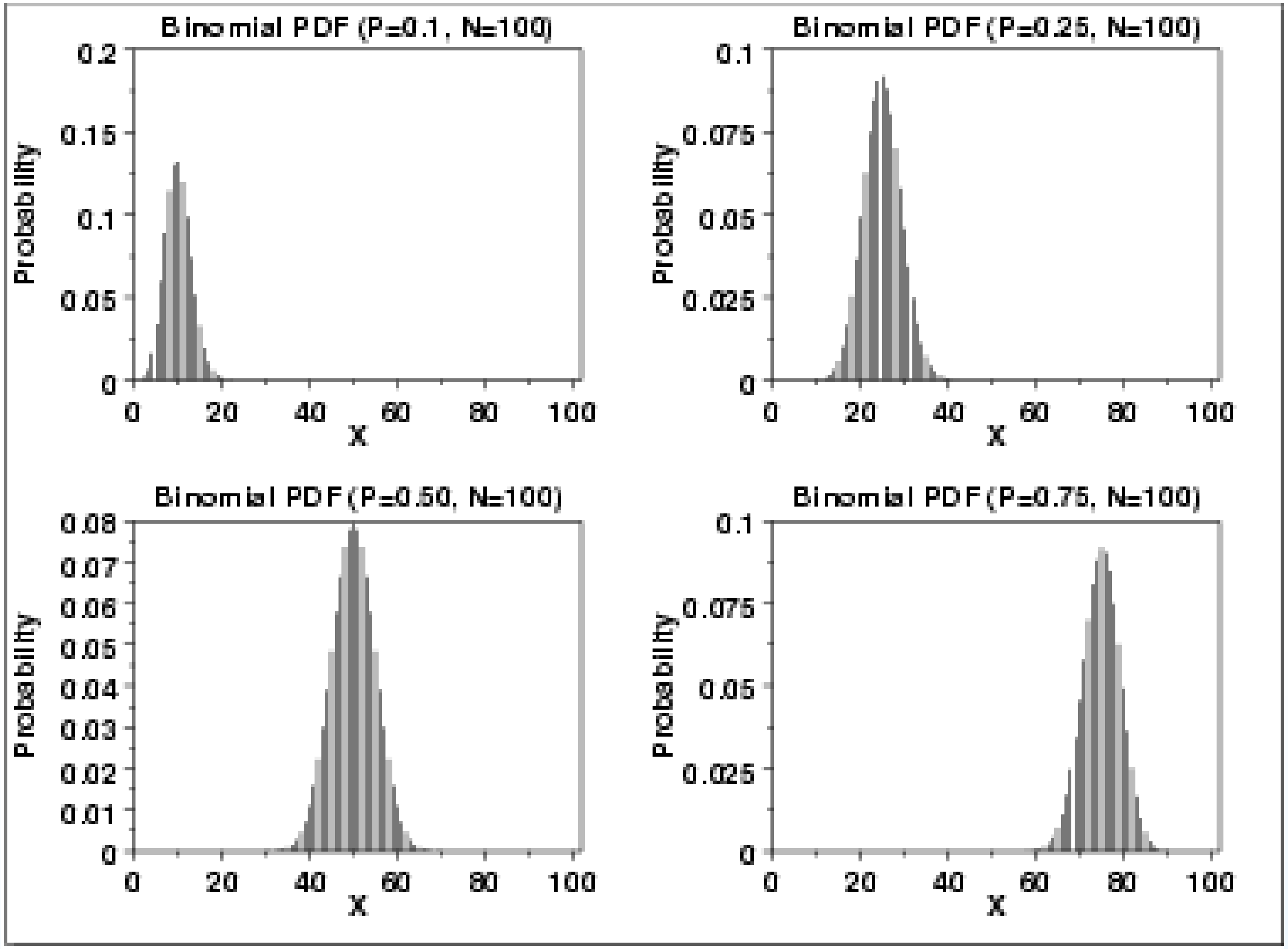
Binomial Distribution

- n draws of a Bernoulli distribution
  - X_i~Bernoulli(p), $X=\sum_{i=1}^{n} X_i$, X~Bin(p, n)
- Random variable X stands for the number of times that experiments are successful.

$$\Pr(X = x) = p_\theta(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

- E[X] = np, Var(X) = np(1-p)

Plots of Binomial Distribution



Poisson Distribution

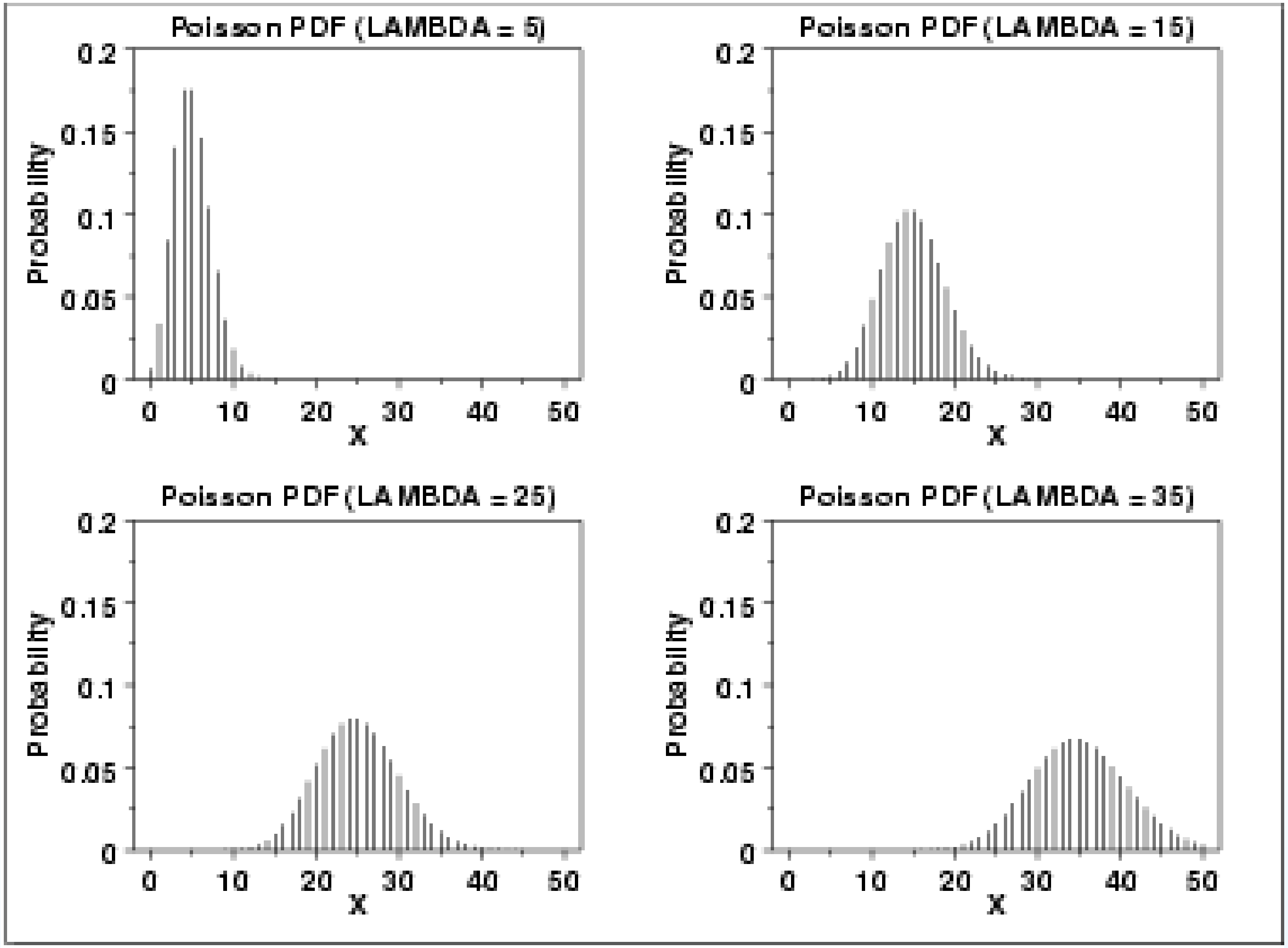
- Coming from Binomial distribution
  - ➢ Fix the expectation $\lambda = np$
  - ➢ Let the number of trials $n \to \infty$

  A Binomial distribution will become a Poisson distribution

$$\Pr(X = x) = p_\theta(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- $E[X] = \lambda$, $Var(X) = \lambda$

Plots of Poisson Distribution



Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$

$$p_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

$$\Pr(a \le X \le b) = \int_a^b p_\theta(x)dx = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$

- $E[X] = \mu$, $Var(X) = \sigma^2$
- If $X_1 \sim N(\mu_1, \sigma^2{}_1)$ and $X_2 \sim N(\mu_2, \sigma^2{}_2)$, for $X = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma^2{}_1 + \sigma^2{}_2)$
- Note that Binomial distributions are Normal (Gaussian)