

# Final Recap (The Good Parts)

IE231 - Lecture Notes - Recap

May 14, 2018

The purpose of this lecture is to give you a complete, quick tour about what happened this semester and what significant points that you should remember. But be warned: These topics, by all means, are not comprehensive.

## Probability

- Probability is the quantification of uncertainty. Sometimes exact probabilities can be calculated (e.g. coin tosses, card draws), sometimes you need to assign probabilities based on assumptions (e.g. outcome of an election, price of bitcoin next month).
- Probabilities cannot be negative, sum of all probabilities in the sample space should equal to one. No event can have a probability greater than 1.
- Probability of two or more events happening together is called a joint probability.
- If several events are disjoint (independent), the probability of all events to happen is the product of their idiosyncratic probabilities to happen  $P(A \cap B) = P(A)P(B)$ .

## Counting

- One of the methods of calculating (discrete) probabilities is to count the sample space. The probability of an event is the relative of its size to the size of the sample space. For example in two coin tosses, the sample space consists of 4 outcomes (i.e. TT, HH, TH, HT) and if we define event A as “one head and one tail in any order”, then the size of our event consists of 2 outcomes (i.e. TH, HT). Therefore, the probability of event A to happen is  $2/4 = 0.5$ .
- You learned three different counting rules; multiplication, permutation and combination.
- Multiplication rule works on repeated experiments with no effect on the sample size. (e.g. coin tosses)
- Permutation rule works on repeated experiments with changing sample spaces (e.g. card draws) and ordering is important.
- Combination rule works when ordering is not important. (e.g. a team of X individuals from a group of Y people)

## Conditional Probability

- Conditional probability is the probability of an event A given the information that another event B happened. For example, probability of “getting two tails” (1/4) is different than “getting two tails given that first coin is a tail” (1/2).
- Conditional probability can be found using the equation  $P(A|B) = \frac{P(A \cap B)}{P(B)}$ . Also,  $P(A|B)P(B) = P(B|A)P(A)$ .
- If A and B are disjoint events  $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$ .
- If there are multiple events  $B_i$  considering event A, then  $P(A) = \sum_i P(A|B_i)P(B_i)$ . (Theorem of total probability)
- If we know event A happened and we want to know the probability of  $B_i$ , then  $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_i P(A|B_i)P(B_i)}$ . (Bayes' Theorem)

## Random Variables and Probability Distributions

- A random variable (say,  $X$ ) is a variable with probabilistic outcome. In other terms we cannot decide its value (decision variable) or it is not a deterministic value (i.e. constant). The outcome of a random variable is usually governed by its underlying distribution.
- A distribution (discrete or continuous) is just a special function.
- What makes a function a distribution is parallel to the axioms of probability. No value given by the function can be less than zero  $f(x) \geq 0$  and total probability (summation for discrete  $\sum_i f(x_i)$ , integral for continuous distribution  $\int_a^b f(x)dx$ ) should be equal to 1.
- Point value probabilities for discrete distributions ( $f(x) = P(X = x)$ ) is called probability mass function and its continuous part ( $f(a < X < b)$ ) is called probability density function. You will see them a lot in the future.
- The sum of probabilities up to a value is called cumulative distribution function for both discrete ( $F(x) = P(X \leq x_k) = \sum_i^k P(X = x_i)$ ) and continuous ( $F(a) = \int_{-\infty}^a f(x)dx$ ) distributions.

## Expectation and Variance

- Expectation is simply a weighted average, where the average is taken by the outcome ( $g(X)$ ) and weights are probabilities. For example, suppose I throw a die and if I roll 5 or 6 I get 15 TL; otherwise I get 0 TL. My expected earnings per roll by intuition is  $15 \cdot 1/3 = 5$  TL. It can also be calculated as  $E[X] = \sum_{i=1}^6 g(i)P(X = i)$  where  $g(5) = g(6) = 15$  and  $P(X = 5) = P(X = 6) = 1/6$ . Then  $0 \cdot 1/6 + 0 \cdot 1/6 + 0 \cdot 1/6 + 0 \cdot 1/6 + 15 \cdot 1/6 + 15 \cdot 1/6 = 15 \cdot 2/6 = 5$ . For continuous distributions it is similar  $\int_{-\infty}^{\infty} g(x)f(x)dx$ .
- Variance is the measure of squared distance of events from the expectation  $V(X) = \sum_i (g(x_i) - E[X])^2 P(X = x_i)$ . It can also be calculated with  $V(X) = E[X^2] - E[X]^2$ .

## Joint, Conditional and Marginal Distributions

- Suppose there are two random variables  $X$  and  $Y$  and we are interested in their joint probabilistic behavior. Their probability mass/density function is defined as  $f(x, y)$ . Joint distribution has the same properties of any probability distribution.
- Marginal distribution is the univariate distribution after the other random variable's effect is "summed out". It can be defined as  $g(x) = \sum_Y f(x, y)$  or  $g(x) = \int^Y f(x, y)dy$ .
- Joint and conditional distributions are similar to joint and conditional probabilities, just that they are functions instead of predefined values.  $f(y|x) = f(x, y)/g(x)$ .

## Some Discrete Distributions

There are some "named" distributions which are useful in application. The most important discrete distributions are **Binomial** and **Poisson**.

- Bernoulli: Binary event, success or fail (e.g. single coin flip). Probability of success is defined with  $p$ .  $E[X] = p$ .
- Binomial: Consists of multiple (say,  $n$ ) bernoulli trials. Probability of  $k$ ;  $k \leq n$  successes are calculated as  $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ .  $E[X] = np$ .
- Multinomial: Multiple probabilities  $p_1, p_2, \dots, p_j$ . Probability of getting  $k_1, k_2, \dots, k_j$  outcomes out of  $n = \sum_i^j k_i$  is calculated with  $P(X = k_1, k_2, \dots, k_j) = \binom{n}{k_1, k_2, \dots, k_j} p_1^{k_1} p_2^{k_2} \dots p_j^{k_j}$ .  $E[X] = \sum_i^k k_i p_i$ .
- Hypergeometric: Multiple binomial from different sources. It can be derived by a combination of counting rules.  $P(X) = \frac{\binom{n}{x_1} \binom{N-n}{x_2}}{\binom{N}{x_1+x_2}}$ .

- Negative Binomial: Very trivial Binomial variation “What is the probability that k-th success occurs at n-th trial?”. We know the last trial should be a success so k-1 trials should occur in any order at n-1 trials.  $P(X) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$ .
- Geometric Distribution: Probability of getting the first success at the n-th trial. Simply  $P(X) = (1-p)^{n-1} p$ .  $E[X] = 1/p$ .
- Poisson: Number of events happening given a time interval. It is a variation of Binomial distribution where n is too high and p is too low. Poisson parameter  $\lambda$  is equivalent to  $np$ .  $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$ .  $\lambda$  parameter can be scaled linearly (e.g. if number of arrivals is poisson with  $\lambda = 2/\text{hour}$  then it is also  $\lambda_h = 1/\text{half-hour}$ ). Therefore sometimes  $\lambda$  is represented with a scaling parameter  $\lambda t$ .

## Some Continuous Distributions

The most important discrete distributions for you to learn are Uniform, Exponential and Normal distributions. It might also be good to hear about gamma, beta and weibull distributions.

- Uniform: Same probability within an interval (say, between a and b).  $f(x) = \frac{1}{b-a}$ ,  $F(x) = \frac{x-a}{b-a}$ .  $E[X] = (b+a)/2$ ,  $V(X) = 1/12(b-a)^2$ .
- Exponential: Closely associated with poisson distribution. Same  $\lambda$  parameter. Exponential distribution measures time between two events (interarrival time), or time before an event happens (e.g. breaking of a light bulb).  $f(x) = \lambda e^{-\lambda x}$ ,  $F(X) = 1 - e^{-\lambda x}$ ,  $E[X] = 1/\lambda$ ,  $V(X) = 1/\lambda^2$ .
- Normal: Most popular distribution. Symmetric. Two parameters location  $\mu$  and scale  $\sigma$ .  $E[X] = \mu$ ,  $V(X) = \sigma^2$ . z-table is used to calculate cumulative probabilities. Convert to standard normal with  $X = \frac{x-\mu}{\sigma}$ .  $\Phi(X)$  means the CDF value of the standard normal distribution.  $\Phi(-X) = 1 - \Phi(X)$  On the graph, x-axis is always the value and area is the cumulative probability given by the table values of the z-table. If a value on the x-axis is also called the quantile.  $\Phi(1.96) = 0.975$  and  $\Phi(-1.96) = 0.025$ . So the area between -1.96 and 1.96 is the 95% of the area (this is important because in statistics you will use these values to calculate confidence interval).

## Conditional Expectation

- Conditional expectation is calculated for conditional distributions ( $E[X|Y = y]$ ).
- They are similar to conditional distribution calculations  $E[X|Y = y] = \sum_x x P(X = x|Y = y)$  for discrete or  $E[X|Y = y] = \int_x x f(x|y) dx$  for continuous.
- $E[E[X|Y]] = \sum_y E[X|Y = y] P(Y = y) = E[X]$ .

## Mini In-Class Exercises

1. In how many ways can you arrange the letters of “SCIENTISTS”?
  - a) Any order?
  - b) Vowels together?
  - c) Vowels separate?
2. In a box there are 15 balls, 5 white 10 black. If I randomly pick 6 balls from the box, what is the probability that it will be 3 white and 3 blacks?
3. There are 18 people; 8 women, 10 men. Suppose you want to form a group of 4 people with at least 1 women or man. How many ways are there?

4. In a marmelade shop people buy strawberry wp 0.5, apricot 0.3 and peach 0.2. Those who buy strawberry likes the marmelade wp 0.8, apricot 0.7 and peach 0.9.
  - a) What is the probability that a random buyer will like the product he/she bought?
  - b) Suppose a buyer did not like her marmelade. What is the probability that the bought apricot marmelade?
5. Ayşe is an actress. She goes to auditions to get roles in movies. She gets an offer from an auditor with probability 0.6. She went to 8 auditions last week.
  - a) What is the probability that she got 4 offers?
  - b) What is the probability that she got at least 2 offers?
6. The cake shop has three kinds of cakes; chocolate cake, strawberry cake and cheesecake. A customer orders chocolate cake with probability 0.3, strawberry cake 0.55 and cheesecake 0.15.
  - a) What is the probability that at least three customers among first 5 customers order chocolate or strawberry cakes?
  - b) What is the probability that the first cheesecake is ordered by the 3rd customer or before?
7. A machine produces 20 items, 12 of which are non-defective. The items are randomly selected without replacement. The sixth selected item is found to be non-defective. What is the probability that this is the third non-defective one?
8. When a Simpsons fan is asked “Which character from Simpsons family is your favorite?”, he/she answers Homer with probability 0.3, Bart w.p. 0.2, Lisa w.p. 0.1, Maggie w.p. 0.2 and Marge w.p. 0.2. In a room of 12 Simpsons fans, what will the probability that there are 6 of them favor Homer, 2 Lisa and 4 Bart?
9. A player throws darts to a special circular target consisting of 4 score regions all centered on the same point. First score region has a radius of  $r$  and second region has a radius of  $2r$ , 3rd region  $3r$  and 4th region  $4r$ . If the player scores at the 1st region he gets 50 points, 2nd region 25, 3rd 10 and 4th 5 points. He has an equal chance to throw the dart within the target (assume probability of missing the target is zero). If he throws 10 darts, what is his expected total score?
10. A bowling player has the probability 0.8 to score a strike at each shot. He makes a bet with his friends If he makes 8 strikes out of 10 he will be given 5 TL, but if he makes fewer than or equal to four strikes he will lose 10 TL.
  - a) What is the probability that he gets the money?
  - b) What is the expected earnings of the player?
  - c) What is the variance of his earnings?
11. People arrive at a concert hall with poisson rate 10 per minute. The concert hall has a capacity of 500 people.
  - a) What is the probability that the concert hall is full in an hour?
  - b) The manager wants to know when the concert hall will be completely full. Give her a time with 90% probability of being true.
12. Time between customer arrivals in a cafe is exponential with the mean value of 6 minutes.
  - a) What is the probability that no customers arrive in 15 minutes?
  - b) What is the interarrival time if the probability of a customer to arrive is 0.9?
  - c) What is the probability that 10 customers arrive in the first hour?
  - d) What is the probability of getting the first customer in 15 minutes if no customer arrived in the first 10 minutes?

13. There are two alarms, each sets off at any random time within 10 and 20 minutes respectively. What is the probability that the second alarm (20 min) sets off later than the first one (10 min)?
14. There is a speed radar between the 15th and 45th km of a 60 km road, randomly placed. In order to avoid a traffic penalty, the speed of a car should be below 60 km/h. The car's max speed is 120 km/h. What is the expected time of finishing the road if the driver wants to keep the risk of getting caught under 25% probability? (*Assume you can instantly change the speed of the car, no acceleration or deceleration*)
15. Meeting durations in a company are normally distributed with mean 1 hour and standard deviation 10 minutes.
  - a) What is the probability that a meeting finishes within 45 minutes?
  - b) What is the duration of a meeting that could have lasted longer with 5% probability?
16. A rice package consists of 750 gr of rice. Though the weight of the package varies with standard deviation of 15 gr.
  - a) What is the probability that the package contains at least 775 grams of rice?
  - b) What is the range of package weight around the mean 90% of the time?
  - c) In order to prevent light weighted packages what should be the mean weight so that 90% of the packages weight 750 gr or more?
17. Three balls are chosen randomly from an urn consisting of 7 green balls and 4 yellow balls. Let  $Y_i$  be the  $i$ th pick, it is 1 if green, 0 otherwise.
  - a) Find the joint probabilities of  $X_1$  and  $X_2$ .
  - b) Find the joint probabilities of  $X_1$ ,  $X_2$  and  $X_3$ .
18. 10% of lightbulbs expire in the first 1000 hours. What is the probability that a lightbulb wouldn't expire in at least 1500 hours? (*Assume exponential distribution*)
19. X and Y have the joint pdf of  $f(x, y) = x + cy^2$ ,  $0 < x < 1$  and  $0 < y < 1$ .
  - a) Find  $c$ .
  - b) Find  $P(X \leq 0.5, Y \leq 0.5)$
20. You are playing the three cups game. There is a marble in one of the cups. You pick a cup at random and if you choose the cup with the marble, you win. Each play costs 5 TL and if you win you earn 15 TL. What is your expected earnings if you play until you win?