STAT381 **Lia Liu**

Chap 2: Probability Models

List of topics: Section 2.1

1. **Experiment: a process of observations or measurements.**
2. **Outcomes: the results obtained from an experiment.**
3. **Sample space(universal set): the set of all possible outcomes of an experiment.**

   Each outcome in a sample space is called an ELEMENT or a SAMPLE POINT.
4. **Event(subset): An EVENT is a subset of a sample space.**
5. **Notation:** $S = \{a, b, c, d, e\}$, **or** $R = \{$**all real numbers**$\}$
6. **Each object(element) counts only once.**

   Ex: $\{1, 2, 2, 2, 3\} = \{1, 2, 3\}$
7. **Special sets:**
   - *Empty set(null set)* **is the set with no elements, written as** Φ.
   - *Universal set* **is the set containing all objects we are considering for a problem. You will always be given the universal set.**
   - *Subset*

     ex: $\{1, 3, 5\}$ **is a subset of** $\{1, 2, 3, 4, 5, 6\}$
   - *Union* **of two sets ("or" relation):** $A \bigcup B =$ **{ elements that are either in A or in B or in both}**
   - *Intersection* **of two sets ("and" relation):** $A \bigcap B =$ **{ elements that belongs to both A and B}**
   - *Complement* **of set A ("not" relation):** A' **or** A^C **contains all elements in the universal set that are not in A. Ex:** $U = \{a, b, c, d, e\}, A = \{b, c, d\}$, **then** $A' = \{a, e\}$
8. **Venn diagram**

   Ex: Complement, union, intersection, $R \bigcap S \bigcap T$
9. **The number of elements of a set is called** *cardinality* **of the set, denoted by** $n(A) =$ **number of elements of A. Ex:** $n\{1, 3, 5, 7\} = 4$

10. **Inclusion-exclusion principle:**

    $n(A \bigcup B) = n(A) + n(B) - n(A \bigcap B)$

    $n(A \bigcup B \bigcup C) = n(A) + n(B) + n(C) - n(A \bigcap B) - n(A \bigcap C) - n(B \bigcap C) + n(A \bigcap B \bigcap C)$

11. **Two sets are** *disjoint or mutually exclusive* **if their intersection is empty.**

    i.e. If $S \bigcap T = \Phi$, **then** S **and** T **are called mutually exclusive.**

    Question 1: Are A **and** A' **always mutually exclusive for any set** A**?**

    Question 2: What becomes the Inclusive-exclusive Principle when S **and** T **are mutually exclusive?**

    Question 3: How many subset does the following sets have?

    $A = \{1, 2, 3\}; B = \{a, b, c, d\}; C = \{1, 2, 3, 4, 5\}$

12. **Discrete sample space**

13. **Continuous sample space**

14. **Def: Probability is a function which maps events of a sample space S into** $[0, 1]$

    P1: $P(A) \geq 0$, **for any event** A **in S;**

    P2: $P(S) = 1$**;**

    P3: If $A_1, A_2, ...$ **is a sequence of mutually exclusive events of S, then** $P(A_1 \bigcup A_2 \bigcup ...) = P(A_1) + P(A_2) + ...$

15. **Complement:** $P(A) = 1 - P(A')$

16. **Inclusive-exclusive principle(again):** $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$

    $P(A \bigcup B \bigcup C) = P(A) + P(B) + P(C) - P(A \bigcap B) - P(A \bigcap C) - P(B \bigcap C) - P(A \bigcap B \bigcap C)$

17. **DeMorgan's rule:**

    $A' \bigcap B' = (A \bigcup B)'$

    $A' \bigcup B' = (A \bigcap B)'$

18. **Venn Diagram**

Section 2.2

List of topics:

1. **Binomial coefficients (will study in detail in 2.4. Just need to know how to compute here.)**
2. **Conditional Probability:** $P(A|B) = \dfrac{P(A \bigcap B)}{P(B)}$
3. **Product rule:**

   $P(A \bigcap B) = P(A|B)P(B) = P(B|A)P(A)$

   $P(A \bigcap B \bigcap C) = P(A)P(B|A)P(C|A \bigcap B)$
4. **Independent: Two events are** *independent* **if**

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

   which means $P(A|B) = P(A)$ **and** $P(B|A) = P(B)$
5. **Tree Diagram**
6. **Partition (Pizza): If**

   1) $A_1, A_2, ..., A_k$ **are mutually exclusive,**

   2) $\bigcup_{i=1}^{k} A_i = S$**;**

   3) $P(A_i) > 0$ **for all** $i = 1, 2, ..., k$.

   Then we say that $A_1, A_2, ..., A_k$ **form a partition.**
7. **The Law of Total Probability: If** $A_1, A_2, ..., A_k$ **form a partition, then for any event** B**,**

   $P(B) = \Sigma_{j=1}^{k} P(B|A_j)P(A_j)$
8. **Smallest partition:** $S = A \bigcup A'$

   $P(B) = P(B|A)P(A) + P(B|A')P(A')$
9. **Bayes' Theorem: If** $A_1, A_2, ..., A_k$ **form a partition, then for any event** B**,**

   $P(A_j|B) = \dfrac{P(B|A_j)P(A_j)}{P(B)}$
10. **Independence vs. Mutually exclusive**

Sections 2.3

List of topics:

1. **Random variable(r.v.): a** *random variable* **is a function from sample space to real numbers.**
2. **Probability distribution function for discrete r.v.(pdf)**

   1) $P(X = k) \geq 0$;

   2) $\Sigma_k P(X = k) = 1$.
3. **How to denote pdf?**

   By table;

   By formula;

   Histogram;

   Relative frequency table.
4. **Expectation of a discrete random variable:**

   $\mu = E(X) = \Sigma_{\textbf{all possible } k} k \cdot f(k)$
5. **Expectation of a function of a discrete random variable:**

   $Eu(X) = \Sigma_{\textbf{all possible } k} u(k) \cdot f(k)$
6. **Expectation is linear:** $E(aX + bY) = aE(X) + bE(Y)$ **for any constants** a, b**, and any random variables** X, Y**.**
7. **Variance of X:**

   $\sigma^2 = V(X) = E(X - \mu)^2$ **(Good for understanding)**

   $\sigma^2 = V(X) = E(X^2) - \mu^2 = (\Sigma k^2 \cdot f(k)) - \mu^2$ **(Good for computation)**
8. **Variance is not linear.**

   i.e. $V(aX + bY) \neq aV(X) + bV(Y)$

   Actually $V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCov(X, Y)$
9. **Standard deviation of** X**:** $\sigma = \sqrt{V(X)}$
10. **Chebychev inequality:**

    $P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \dfrac{1}{k^2}$

Sections 2.4

List of topics: Binomial Distribution

1. **Binomial expansion:** $(x+y)^n$
2. **Binomial coefficients:** $C(n,k)$

   1) By Yang's triangle

   2) By combination: $C(n,k) = \dfrac{P(n,k)}{k!} = \dfrac{n!}{k!(n-k)!}$
3. **Bernoulli trials:**
4. **Binomial distribution: check 3 conditions: exactly 2 outcomes for each trial; independent trials; probability of heads stays the same for all trials.**

   Model: Toss a coin n times.

   $P(H) = p, P(T) = q, q = 1-p$**.**

   X= number of heads.
5. **p.d.f.** $P(X=k) = C(n,k)p^k q^{n-k},\ k = 0,1,2,...,n$
6. **Expected value** $E(X) = np$; **or** $\mu = np$

   Variance $V(X) = npq$; **or** $\sigma^2 = npq$

   Standard deviation $\sigma = \sqrt{npq}$
7. **Geometric distribution: Toss a coin until the 1st H.**

   pdf: $P(X=k) = q^{k-1}p, k = 1,2,3,...$

   Mean: $E(X) = 1/p$**. (How to find out? Using term-by-term differentiation formula.)**

   Variance: $V(X) = q/p^2$**. Hint:Find** $E[X(X-1)]$ **first by using term-by-term differentiation formula twice.**
8. **Negative Binomial distribution: Toss a coin until rth heads**
9. **Hypergeometric distribution: without replacement**

Sections 2.5

List of topics: Poisson Distribution

1. **Model: Counting number of occurrences. (e.g. number of cars passing an intersection for a duration of time, number of calls going through a switchboard, number of login requests on a website during a time period, number of flaws on a piece of fabric).**

2. **Assumptions: Use number of occurrences during a time interval as an example:**

   1. The probability of exactly one occurrence in a short time interval is roughly proportional to the interval length.

   2. It is unlikely to have two or more changes in a short time;

   3. The number of occurrences in nonoverlapping intervals are mutually independent.

3. **pdf:** $P(X=k)=\dfrac{e^{-\lambda}\lambda^k}{k!},\ k=0,1,2,...$

4. **mean and variance are both** λ

5. **Closely related to Exponential distribution. Will discuss in section 3.3.**

If $g(x,w)$ **is the probability of** x **occurrences in an interval of length** w**, then the forts two assumptions becomes**

1. $g(1,h)=\lambda h+o(h)$

2. $\sum_{x=2}^{\infty} g(x,h)=o(h)$.

Then $g(0,w+h)=g(0,w)[1-\lambda h-o(h)]$**, so**

$$\frac{g(0,w+h)-g(0,w)}{h}=-\lambda g(0,w)-\frac{o(h)g(0,w)}{h}$$

Take the limit as $h\to 0$**,** $\frac{d}{dw}g(0,w)=-\lambda g(0,w)$**.**

The solution of this diff equation is $g(0,w)=ce^{-\lambda w}$

But $g(0,0)=1$**,** $g(0,w)=e^{-\lambda w}$**.**

Section 2.6 Dist of 2 discrete random variables

List of topics:

1. **Joint pdf for two discrete random variables:**

   $f_{X_1,X_2}(x,y) = P(X_1 = x, X_2 = y)$

   Note: This f **is a probability, hence** $0 \le f \le 1$

2. **CDF of** X_1 **and** X_2**:**

   $F_{X_1,X_2}(x,y) = P[(X_1 \le x) \bigcap (X_2 \le y)]$

3. **Marginal pdfs: The pdf of one variable. Sum up against the other variable:**

   $f_X(x) = \Sigma_{ally} f(x,y)$; $f_Y(y) = \Sigma_{allx} f(x,y)$;

4. **Expectations**

5. $E(u(X_1, X_2))$

6. **Conditional pdf is the ratio of joint pdf to marginal pdf.**

   $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$ **for** y **in the support of** Y**.**

   $f_{Y|X}(y|x)$ **is Similarly defined.**

7. **Conditional expectation:**

   $E(u(Y)|X = x) = \sum u(y) f_{Y|X}(y|x)$ **(which is a function of** x**.)**

   Hence $E(Y|X)$ **is a random variable of** X**.**

8. **Covariance of X and Y:**

   $cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$ **(easier to understand)**

   $cov(X,Y) = E(XY) - \mu_X\mu_Y$ **(easier for calculation)**

9. **Correlation coefficient** $\rho = \dfrac{cov(X,Y)}{\sigma_X\sigma_Y}$

10. **Example: If X and Y have joint pdf:**

		X			
		0	**1**	**2**	f_Y
Y	**0**	**1/18**	**4/18**	**6/18**	**11/18**
	1	**3/18**	**3/18**	**1/18**	**7/18**
f_X		**4/18**	**7/18**	**7/18**	

   Note the marginal distributions are written on the "margins" of the table.

   Conditional distribution of X given $Y=1$ is

   $f_{X|Y}(0|y=1) = \dfrac{f(0,1)}{f_Y(1)} = 3/7;$

   $f_{X|Y}(1|y=1) = \dfrac{f(1,1)}{f_Y(1)} = 3/7;$

   $f_{X|Y}(2|y=1) = \dfrac{f(2,1)}{f_Y(1)} = 1/7.$

   To find $E(X)$, you can use the marginal distribution of X,

   $E(X) = 0 + 1*7/18 + 2*7/18 = 21/18;$

   or the joint pdf:

   $E(X) = \Sigma_{x=0}^{2}\Sigma_{y=0}^{1} xf(x,y)$

   To find $E(XY)$ you have to use the joint pdf:

   $E(XY) = \Sigma_{x=0}^{2}\Sigma_{y=0}^{1} xyf(x,y) = 0+1*1*3/18+1*2*1/18 = 2/18$

11. **Equivalent conditions: X and Y are independent iff**

   (1) By pdf: $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ **(for each pair of x and y.)**

   (2) By CDF: $F(x,y) = F_X(x)F_Y(y)$ **for all** (x,y)

   (3) By MGF: $M(t_1,t_2) = M(t_1,0)M(0,t_2)$

12. **How can you tell two r.v. X and Y are independent by joint density?**

    Check each entry in the joint density table is the product of two marginals.

13. **What happens to covariance(and correlation coefficient) if X and Y are independent?**

    If X and Y are independent, then $Cov(X,Y) = 0$, hence $\rho = 0$.

    However, if $Cov(X,Y) = 0$, or $\rho = 0$, X and Y may not be independent.

Facts:

- $-1 \leq \rho \leq 1$
- $\rho = \pm 1$ **implies"perfect linear relationship between X and Y."**
- $\rho = 0$ **implies "No linear relationship between X and Y (but can be quadratic, for example)."**
- $\rho > 0$ **means X and Y are positively related (If X increases, so is Y).**
- $\rho < 0$ **means X and Y are negatively related (If X increases, Y decreases).**
- **If X and Y are independent, then $Cov(X,Y) = 0$, hence $\rho = 0$.**

  However, if $Cov(X,Y) = 0$, or $\rho = 0$, X and Y may not be independent.

  Ex: 2.5.10.
- **Special case: If X and Y are normal (covered later), then $Cov(X,Y) = 0$, or $\rho = 0$ implies "X and Y are independent."**