# Homework 2

S1201, Summer 2022

100 points

Due July 20, 23:59PM EST

**Notes:** Please read the following instructions.

(i) Please submit a **single** PDF file on courseworks. You may scan your written solutions or directly write the solutions on your tablet.

(ii) Late submission will lead to some penalty: 20 points off for within 24 hours late, 40 points off for over 24 hours late, all points off for submission after the solutions are posted on courseworks.

(iii) You can discuss the problems with others. But a direct plagiarism will lead to zero point for this assignment, and this will be reported to the university.

(iv) There is one extra problem with 7 bonus points, which is optional. Therefore all points add up to 107, but the maximum score of this homework is 100.

1. $\mathbf{2 \times 7 = 14}$ **points** TRUE/FALSE questions. No explanations are needed.

   (a) The cumulative distribution function, $F(x)$, is never greater than 1. T

   (b) If for some events $A, B, C$, we have $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.4$ and $\mathbb{P}(C) = 0.5$, then it is possible for $\mathbb{P}(A \cup B \cup C)$ to equal 1.2. F

   (c) If 20% of all students at Columbia are taking a statistics course and 10% are taking a history course, then 30% are taking a statistics **or** history course. F

   (d) For a continuous variable $X$, rather than $F(x) = \mathbb{P}(X \leq x)$, its cdf can be also be defined as $F(x) = \mathbb{P}(X < x)$. T

   (e) If $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C)$, then events $A$, $B$ and $C$ are mutually exclusive. F

   (f) The random variable $X$ is a function, which maps an outcome to a number. T

   (g) The permutation number $P_{k,n} = k! \times$ the combination number $\binom{n}{k}$ T

2. **2 × 4 = 8 points** Suppose $\mathbb{P}(A) = 0.6$ and $\mathbb{P}(B) = 0.59$, where $A$ and $B$ are some random events.

   (a) What are the maximum and the minimum possible values for $P(A \cap B)$? Draw Venn diagrams for each situation.

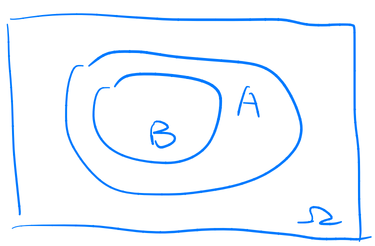
   (b) What are the maximum and the minimum possible values for $P(A \cup B)$? Draw Venn diagrams for each situation.

   (c) Is it possible for $A$ and $B$ to be mutually exclusive? If yes, provide an example; if no, explain why.

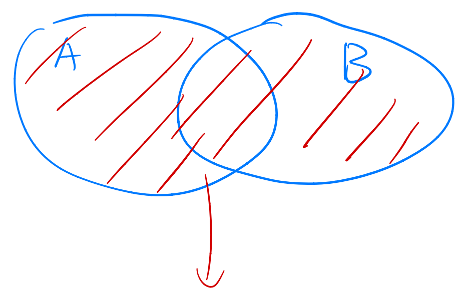
   (d) Is it possible for $A$ and $B$ to be independent? If yes, provide an example; if no, explain why.

Solution: Let's consider (b) first.

(b) $0.6 \le \mathbb{P}(A) \le \mathbb{P}(A \cup B) \le 1.$




$\mathbb{P}(A \cup B) = 0.6$




$\Omega = A \cup B.$

(a) $$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$$
$$= 1.19 - \mathbb{P}(A \cup B)$$
$$\in [0.19, 0.59]$$

(c) No. Otherwise $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 1.19 > 1.$

(d) Yes.

Example: Drawing a number from 1, 2, 3, ..., 100 twice (with replacement).

$A = \{$ 1st draw $= 1, 2, \ldots,$ or $60 \}.$

$B = \{$ 2nd draw $= 1, 2, \ldots,$ or $59 \}$ $\square$.

# Homework 2

S1201, Summer 2022

100 points — Due July 20, 23:59PM EST

3. **3 × 2 = 6 points** Answer the following two questions.

(a) $f(x) = \begin{cases} ae^{-x}, & 0 \le x \le 1, \\ \frac{1}{2x^2}, & x > 1, \\ 0, & \text{elsewhere}, \end{cases}$. What's the value of $a$ that makes $f(x)$ a legitimate pdf?

(b) With the value of $a$ you derive in part (a), write out the corresponding cdf function.

Solution: (a) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

$$\Rightarrow a\int_0^1 e^{-x}\,dx + \int_1^{+\infty} \frac{1}{2x^2}\,dx = 1$$

$$\Rightarrow a(1-e^{-1}) - \frac{1}{2x}\Big|_{x=1}^{x=+\infty} = 1$$

$$\Rightarrow a = \frac{1}{2(1-e^{-1})}.$$

(b) $F(x) = \mathbb{P}(X \le x)$

$$= \int_{-\infty}^{x} f(t)\,dt$$

$$= \begin{cases} \int_0^x ae^{-t}\,dt, & \text{if } 0 \le x \le 1 \\ \int_0^1 ae^{-t}\,dt + \int_1^x \frac{1}{2t^2}\,dt, & \text{if } x > 1 \\ 0, & \text{if } x < 0 \end{cases}$$

$$= \begin{cases} \frac{1-e^{-x}}{2(1-e^{-1})}, & \text{if } 0 \le x \le 1 \\ 1 - \frac{1}{2x}, & \text{if } x > 1 \\ 0, & \text{if } x < 0. \end{cases}$$

□.

4. **$3 \times 2 = 6$ points** Answer the following two questions.

(a) $p(x) = \begin{cases} ax^2, & x = 1, 2, 3, \\ 0, & \text{elsewhere}, \end{cases}$. What's the value of $a$ that makes $p(x)$ a legitimate ~~pdf~~? pmf.

(b) With the value of $a$ you derive in part (a), write out the pmf table.

Solution: (a) Let $p(1) + p(2) + p(3) = 1$ :

$$\Rightarrow \quad a(1 + 4 + 9) = 1$$

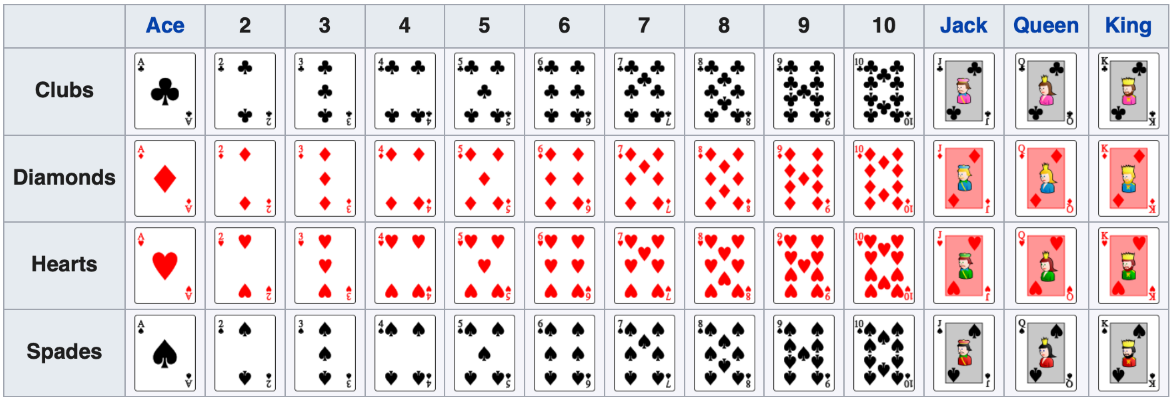
$$\Rightarrow \quad a = \frac{1}{14}$$

(b)

| $x$ | 1 | 2 | 3 |
|---|---|---|---|
| $p(x)$ | $\frac{1}{14}$ | $\frac{4}{14}$ | $\frac{9}{14}$ |

$\square$

5. **3 × 4 = 12 points** 8 cards are selected at random from a 52-card deck.




(a) What's the probability of getting four 8's and four Queens?

(b) What's the probability of getting a hand consisting of all four suits (i.e. have at least one card from each of four suits)?

(c) What is the probability of getting a hand of 2 × ♠-2 × ♡-2 × ♢-2 × ♣? event A

(d) What is the probability of getting a hand with distribution like "4-1-1-2" (e.g.: 4 × ♠-1 × ♡-1 × ♢-2 × ♣, 4 × ♡-1 × ♣-1 × ♢-2 × ♠ etc.) event B

Solution: Combination problem.

(a) $N = N(\Omega) = \binom{52}{8} = 752538150$

$N(\text{getting four 8's \& four Queens}) = \binom{4}{4}\binom{4}{4} = 1.$

$\Rightarrow P(\text{getting four 8's \& four Queens}) = \frac{1}{\binom{52}{8}} \approx 1.33 \times 10^{-9}$

(b) Case 1: 2-2-2-2

very challenging

contains $\left[\binom{13}{2}\right]^4 = 37015056$ outcomes $a_1$

Case 2: 3-2-2-1

contains $\binom{4}{2}\binom{2}{1}\cdot\binom{13}{3}\cdot\left[\binom{13}{2}\right]^2\cdot\binom{13}{1} = 74030112$ outcomes $a_2$

Case 3: 3-3-1-1.

contains $\binom{4}{2}\left[\binom{13}{3}\right]^2\cdot\left[\binom{13}{1}\right]^2 = 8294 1144$ outcomes $a_3$

Case 4: 4-2-1-1

contains $\binom{4}{1}\binom{3}{1}\binom{13}{4}\binom{13}{2}\left[\binom{13}{1}\right]^2 = 113101560$ outcomes $a_4$

Case 5: $5-1-1-1$ $a_5$

Contains $\binom{4}{1}\binom{13}{5}\cdot\left[\binom{13}{1}\right]^3 = 1131015\underline{6}$ outcomes

$\Rightarrow P(4 \text{ suits}) = \dfrac{a_1+a_2+a_3+a_4+a_5}{N} = 0.4081$ $\boxed{\text{Can be wrong.}}$

(c) $N(A) = \left[\binom{13}{2}\right]^4 = 37015056$ (case 4 in (b))

$\Rightarrow P(A) = \dfrac{N(A)}{N} = 0.0492$

(d) If we know the suits of "4-1-1-2", then the number of ~~outcomes~~ would be $\binom{13}{4}\cdot\left[\binom{13}{1}\right]^2\cdot\binom{13}{2}$ by product rule ( step 1: choose 4 from some specific suit, step 2: choose 1 from some specific suit, ....)

The number of ways to arrange 4 suits in "4-1-1-2"

$= \binom{4}{1}\times\binom{3}{2} = 12$ (by product rule)

Step 1: decide which suit has 4 cards

step 2: decide which 2 suits have 1 card.

$\Rightarrow N(B) = \binom{13}{4}\left[\binom{13}{1}\right]^2\times\binom{13}{2}\times 12 = 113101560$

$\Rightarrow P(B) = \dfrac{N(B)}{N} = 0.1503.$ $\square$.

6. **$3 \times 3 = 9$ points** Prove the following conclusions.

   (a) Suppose $\mathbb{P}(B) > 0$. Then events $A$ and $B$ are independent (here we mean the first definition $\mathbb{P}(A|B) = \mathbb{P}(A)$) if and only if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$

   (b) Suppose $0 < \mathbb{P}(B) < 1$. Then events $A$ and $B$ are independent if and only if $A^c$ and $B^c$ are independent;

   (c) If two independent events $A$ and $B$ satisfy $0 < \mathbb{P}(A) < 1$ and $0 < \mathbb{P}(B) < 1$, then $A$ and $B$ cannot be mutually exclusive.

Notes: As long as you show one direction, "if" or "only if", you'll get full marks.

Solution: (a) $\mathbb{P}(A|B) = \mathbb{P}(A) \iff \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A) \iff \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

(b) ① $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Longrightarrow \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c)$ :

In fact: 
$$\begin{aligned}\mathbb{P}(A^c \cap B^c) &= 1 - \mathbb{P}((A^c \cap B^c)^c) \\ &= 1 - \mathbb{P}(A \cup B) \\ &= 1 - (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)) \\ &= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B) \\ &= (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)). \\ &= \mathbb{P}(A^c)\mathbb{P}(B^c) \quad \checkmark\end{aligned}$$

② $\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c) \Longrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

No need to prove (because $(A^c)^c = A$, $(B^c)^c = B$, then just use part ①).

(c) If $A$ & $B$ are mutually exclusive, then:

$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. ⊛

But $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. that means

$\mathbb{P}(A \cup B) = 0$

But by ⊛, $\mathbb{P}(A \cup B) > 0 \Rightarrow$ contradiction!

So $A$ & $B$ cannot be mutually exclusive $\square$.

7. **10 points** About 30% of human twins are identical, and the rest are fraternal. Identical twins are necessarily the same sex, half are males and the other half are females. One-quarter of fraternal twins are both male, one-quarter both female, and one-half are mixes: one male, one female. You have just become a parent of twins and are told they are both girls. Given this information, what is the probability that they are identical?

ID FR M F MF

Solution: We know that $\mathbb{P}(ID) = 0.3$, $\mathbb{P}(FR) = 0.7$

$\mathbb{P}(M \mid ID) = \mathbb{P}(F \mid ID) = 0.5$

$\mathbb{P}(M \mid FR) = \mathbb{P}(F \mid FR) = 0.25$

$\mathbb{P}(MF \mid FR) = 0.5$.

Goal: Calculate $\mathbb{P}(ID \mid F) = ?$

ID & FR are a partition.

By Bayes' Theorem:

$$\mathbb{P}(ID \mid F) = \frac{\mathbb{P}(F \mid ID)\,\mathbb{P}(ID)}{\mathbb{P}(F \mid ID)\,\mathbb{P}(ID) + \mathbb{P}(F \mid FR)\,\mathbb{P}(FR)}$$

$$= \frac{0.5 \times 0.3}{0.5 \times 0.3 + 0.25 \times 0.7}$$

$$\approx 0.4615 \qquad \square.$$

8. **5+5 = 10 points** Suppose that it is known that 20% of students in a certain university live off campus. If we randomly sample 15 students from this university, what is the probability that:

   (a) Exactly 5 live off campus?

   (b) At least one lives off campus?

   Hint: Recall the example of binomial distribution in slides of lecture ~~8~~. 7

Solution: Denote the number of students among these 15 living off campus as $X$.

Then $X \sim \text{Bin}(n=15, p=0.2)$

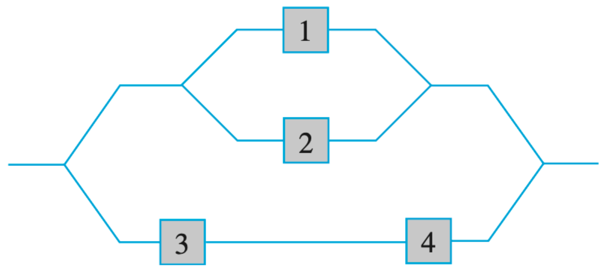
(a) $\mathbb{P}(X=5) = \binom{15}{5} \times 0.2^5 \times 0.8^{10} \approx 0.1032$

(b) $\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0)$

$= 1 - 0.8^{15}$

$\approx 0.9648$ $\square$

9. **6 points** Consider the system of components connected as in the accompanying picture. Components 1 and 2 are connected in parallel, so that subsystem works iff either 1 or 2 works; since 3 and 4 are connected in series, that subsystem works iff both 3 and 4 work. If components work independently of one another and $\mathbb{P}(\underbrace{\{\text{component i works}\}}_{C_i}) = 0.9$ for $i = 1, 2$ and $= 0.8$ for $i = 3, 4$, calculate $\mathbb{P}(\{\text{the system works}\})$.




Solution: $\mathbb{P}(\{\text{system works}\})$

$$= \mathbb{P}((C_1 \cup C_2) \cup (C_3 \cap C_4))$$

$$= \mathbb{P}(C_1 \cup C_2) + \mathbb{P}(C_3 \cap C_4) - \mathbb{P}((C_1 \cup C_2) \cap (C_3 \cap C_4)),$$

where $\mathbb{P}(C_1 \cup C_2) = \mathbb{P}(C_1) + \mathbb{P}(C_2) - \mathbb{P}(C_1 \cap C_2)$

$$= \mathbb{P}(C_1) + \mathbb{P}(C_2) - \mathbb{P}(C_1) \cdot \mathbb{P}(C_2)$$

$$= 0.9 + 0.9 - 0.9^2$$

$$= 0.99$$

$$\mathbb{P}(C_3 \cap C_4) = \mathbb{P}(C_3)\mathbb{P}(C_4) = 0.8^2 = 0.64$$

$$\mathbb{P}((C_1 \cup C_2) \cap (C_3 \cap C_4)) = \mathbb{P}(C_1 \cup C_2) \times \mathbb{P}(C_3 \cap C_4)$$

$$= 0.99 \times 0.64$$

$$= 0.6336$$

$$\Rightarrow \mathbb{P}(\{\text{system works}\}) = 0.99 + 0.64 - 0.6336$$

$$= 0.9964. \quad \square$$

10. **$2 \times 4 + 5 = 13$ points** The accompanying table categorizing each student in a sample according to gender and eye color :

| | Eye Color | | | | |
|---|---|---|---|---|---|
| **Gender** | **Blue** | **Brown** | **Green** | **Hazel** | **Total** |
| **Male** | 370 | 352 | 198 | 187 | 1107 |
| **Female** | 359 | 290 | 110 | 160 | 919 |
| **Total** | 729 | 642 | 308 | 347 | 2026 |

Suppose that one of these 2026 students is randomly selected. Let F denote the event that the selected individual is a female, and A, B, C, and D represent the events that he or she has blue, brown, green, and hazel eyes, respectively.

(a) Calculate both $\mathbb{P}(F)$ and $\mathbb{P}(C)$

(b) Calculate $\mathbb{P}(F \cap C)$. Are the events F and C independent? Why or why not?

(c) If the selected individual has green eyes, what is the probability that he or she is a female?

(d) If the selected individual is female, what is the probability that she has green eyes?

(e) (5 points) What is the "conditional distribution" of eye color for females (i.e., $P(A|F)$, $P(B|F)$, $P(C|F)$, and $P(D|F)$)? And what is it for males (i.e., $P(A|M)$, $P(B|M)$, $P(C|M)$, and $P(D|M)$)? Compare the two distributions.

Solution: (a) $\mathbb{P}(F) = \frac{919}{2026} \approx 0.4536$ , $\mathbb{P}(C) = \frac{308}{2026} \approx 0.1520$

(b) $\mathbb{P}(F \cap C) = \frac{110}{2026} = 0.0543 \neq \mathbb{P}(F) \times \mathbb{P}(C)$

$\Rightarrow F \not\perp C$

(c) $\mathbb{P}(F \mid C) = \frac{\mathbb{P}(F \cap C)}{\mathbb{P}(C)} = \frac{0.0543}{0.1520} \approx 0.3572$

(d) $\mathbb{P}(C \mid F) = \frac{\mathbb{P}(F \cap C)}{\mathbb{P}(F)} = \frac{0.0543}{0.4536} \approx 0.1197$

(e) ① $\mathbb{P}(A \mid F) = 0.3906$ , $\mathbb{P}(B \mid F) = 0.3156$

$\mathbb{P}(C \mid F) = 0.1197$ , $\mathbb{P}(D \mid F) = 0.1741$

② $\mathbb{P}(A \mid M) = 0.3342$ , $\mathbb{P}(B \mid M) = 0.3180$

$\mathbb{P}(C \mid M) = 0.1789$ , $\mathbb{P}(D \mid M) = 0.1689$.

Comparison: Can say anything reasonable. E.g., female population has a larger proportion of blue eyes than male population. □



11. **3 + 3 = 6 points** Airlines sometimes overbook flights. Suppose that for a plane with 50 seats, 55 passengers have tickets. Define the random variable $Y$ as the number of ticketed passengers who actually show up for the flight. The probability mass function of $Y$ appears in the accompanying table.

| $y$ | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p(y)$ | .05 | .10 | .12 | .14 | .25 | .17 | .06 | .05 | .03 | .02 | .01 |

(a) What is the probability that **not** all ticketed passengers who show up can be accommodated?

(b) If you are the first person on the standby list (which means you will be the first one to get on the plane if there are any seats available after all ticketed passengers have been accommodated), what is the probability that you will be able to take the flight? What is this probability if you are the third person on the standby list?

Solution: (a) The probability $= \mathbb{P}(Y > 50)$

$= P(51) + P(52) + P(53) + P(54) + P(55)$

$= 0.17$

(b) ① If I'm the 1st person:

The prob that I can take the flight

$= \mathbb{P}(Y \leq 49)$

$= P(45) + P(46) + P(47) + P(48) + P(49)$

$= 0.66$

② If I'm the 3rd person:

The prob that I can take the flight

$= \mathbb{P}(Y \leq 47)$

$= P(45) + P(46) + P(47)$

$= 0.27$. □

12. **(Extra problem, not required) $3+2+2=7$ bonus points** Ignore this

~~Only 1 in 1000 adults is afflicted with a rare disease for which a diagnostic test has been developed.~~ Assume that for such a test, 98% of patients with disease obtain positive results and 97% of patients without disease obtain negative results.

(a) Assume that (100a)% of the population has the disease, where $0 \le a \le 1$. Pick one person at random. The lab test shows positive result. What is the chance that the person really has the disease? Hint: the answer should be a function of $a$.

(b) What condition should $a$ satisfy for the probability in part (a) to be over 1%? What condition should $a$ satisfy for the probability in part (a) to be over 50%?

(c) Draw a plot of the probability in part (a) versus the value of $a$. (Do not need to be very accurate, but the shape should be correct)

Solution: (a) Let $D = \{\text{has the disease}\}$.

$P = \{\text{positive result}\}$.

We know: $\mathbb{P}(D) = a$, $\mathbb{P}(P \mid D) = 0.98$

$\mathbb{P}(P^c \mid D^c) = 0.97$

By Bayes' theorem: $\mathbb{P}(D \mid P) = \dfrac{\mathbb{P}(P\mid D)\cdot \mathbb{P}(D)}{\mathbb{P}(P\mid D)\mathbb{P}(D) + \mathbb{P}(P\mid D^c)\mathbb{P}(D^c)}$

$$= \frac{0.98 \times a}{0.98 \times a + 0.03 \times (1-a)}$$

$$= \frac{0.98a}{0.95a + 0.03}$$

(b) ① Let $\mathbb{P}(D \mid P) \geq 1\% \Rightarrow \dfrac{0.98a}{0.95a + 0.03} \geq 0.01$

$\Rightarrow a \geq 3.09 \times 10^{-4}$

② Let $\mathbb{P}(D \mid P) \geq 50\% \Rightarrow \dfrac{0.98a}{0.95a + 0.03} \geq 0.5$

$\Rightarrow a \geq 0.0297$.

(c)