Week May 21, 2018

Kaiyan Shi

May 26, 2018

Monday

Statistics

work with Iris

2.20

Define the following events: $X =$ Second ball drawn is white, $Z =$ First ball drawn is white.

$$P(X|Z) = \frac{P(XZ)}{P(Z)} = \frac{1/2}{P(Z)}.$$

Then since XZ represents both balls drawn are white, $P(XZ) = \frac{1}{2}$, because the "uncertain" ball has probability $\frac{1}{2}$ of being white.

There are two ways for Z to occur, one where the first ball chosen is the "certainly" white ball, and the the second is where the first ball chosen is the uncertainly white ball. Call these events Y and Y^c, respectively. Y and Y^c partition our sample space since one of these events will always happen, but never at the same time. So,

$$\begin{aligned} P(Z) &= P(Z|Y)P(Y) + P(Z|Y^c)P(Y^c) \\ &= \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \\ &= \frac{3}{4}. \end{aligned}$$

Hence,

$$P(X|Z) = \frac{1/2}{3/4} = \frac{2}{3}$$

3.2

$P(A) = \frac{1}{3}, P(B) = \frac{1}{4}, P(C) = \frac{1}{5}$

a

$P(ABC) = P((A \cap B) \cap C) = P(A)P(B)P(C) = \frac{1}{3} \cdot \frac{1}{4} \cdot \frac{1}{5} = \frac{1}{60}$

b

$$\begin{aligned} P(A \cup B \cup C) &= P((A \cup B) \cup C) \\ &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\ &= P(A) + P(B) - P(A)P(B) + P(C) - (P(A) + P(B) - P(A)P(B))P(C) \\ &= P(A) + P(B) + P(C) - P(A)P(B) - P(B)P(C) - P(A)P(C) + P(A)P(B)P(C) \\ &= \frac{3}{5} \end{aligned}$$

c

$P(ABC|C) = \frac{1/60}{1/5} = \frac{1}{12}$

d

$P(B|AC) = \frac{P(ABC)}{P(AC)} = \frac{P(A)P(B)P(C)}{P(A)P(C)} = P(B) = \frac{1}{4}$

e

$$\begin{aligned} P(\text{at most one of the three events occur}) &= P(A^c B^c C^c) + P(AB^c C^c) + P(A^c B C^c) + P(A^c B^c C) \\ &= \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} + \frac{1}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} + \frac{2}{3} \cdot \frac{1}{4} \cdot \frac{4}{5} + \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{1}{5} \\ &= \frac{5}{6} \end{aligned}$$

4.14

Since the three variables are independent, the joint probabilities are equal to the product of their respective marginal probabilities. So for any $i, j, k \in \{1, 2\}$,

$$P(A = i, B = j, C = k) = (\frac{1}{2})^3 = \frac{1}{8}.$$

And since there are eight different possible events, the sum of these probabiliits equal one.

4.19

a

$$\begin{aligned} P(X = 1) &= \frac{1}{8} + \frac{1}{4} = \frac{3}{8} \\ P(X = 2) &= \frac{1}{8} + \frac{1}{2} = \frac{5}{8} \\ P(Y = 1) &= \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \\ P(Y = 2) &= \frac{1}{4} + \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

b

X and Y are not independent. For independence to occur, $P(X = x) = P(X = x|Y = y)$ must be true for all x and y. This certainly does not occur when $X = 1, Y = 1$, since $P(X = 1) = \frac{3}{8} \neq \frac{1}{2} = P(X = 1|Y = 1)$.

c

$P(XY \leq 3) = 1 - P(X = 2, Y = 2) = 1 - \frac{1}{2} = \frac{1}{2}$

d

$P(X + Y > 2) = 1 - P(X = 1, y = 1) = 1 - \frac{1}{8}$

Cyptography

2.3

Disprove the direction: An encryption scheme with message space $\mathcal{M}$ is perfectly secret $\Rightarrow$ For every probability distribution over $\mathcal{M}$ and every $c_0, c_1 \in C$, $\Pr[C = c_0] = \Pr[C = c_1]$.

DisPf/ Consider one-time pad scheme with cipertext space $\mathcal{C} = \{\{0,1\}^l, 2\}$.
As $\Pr[C = 2] = 0$, and one-time pad is perfectly secrete, the modified scheme also satisfies the first definition of perfect security.
But as $\Pr[C = 2] = 0$, and $\Pr[C = 1] \neq 0$, we have $\Pr[C = 2] \neq \Pr[C = 1]$.

2.6

a

It is not a perfectly secure scheme.
Consider $c = 0, m_1 = 0$ and $m_2 = 1$.
$\Pr[\text{Enc}_k(0) = 0] = \Pr[k + 0 \bmod 5 = 0] = \Pr[k \bmod 5 = 0] = \Pr[k = 0] + \Pr[k = 1] = \frac{1}{3}$.
$\Pr[\text{Enc}_k(1) = 0] = \Pr[k + 1 \bmod 5 = 0] = \Pr[k = 4] = \frac{1}{6}$.
Therefore, $\Pr[\text{Enc}_k(m_1) = c] \neq \Pr[\text{Enc}_k(m_2) = c]$

b

The scheme is perfectly secrete.
Consider $c \in \{0,1\}^{l-1}||0$
$\Pr[\text{Enc}_k(m_1) = c] = \frac{1}{2^{l-1}}$
$\Pr[\text{Enc}_k(m_2) = c] = \frac{1}{2^{l-1}}$
Consider $c \in \{0,1\}^{l-1}||1$
$\Pr[\text{Enc}_k(m_1) = c] = 0$
$\Pr[\text{Enc}_k(m_2) = c] = 0$
Therefore, $\forall c, m_1, m_2$, $\Pr[\text{Enc}_k(m_1) = c] \neq \Pr[\text{Enc}_k(m_2) = c]$.

2.7

The modified one-time pad scheme is not perfectly secrete.
Consider $c = m, m_1 = m, m_2 = m' \neq m$.
$\Pr[\text{Enc}_k(m) = m] = \Pr[m \oplus k = m] = \Pr[k = 0^l] = 0$

$\Pr[\text{Enc}_k(m') = m] = \frac{1}{2^l - 1}$
Therefore, $\Pr[\text{Enc}_k(m_1) = c] \neq \Pr[\text{Enc}_k(m_2) = c]$.

2.13

a

No encryption scheme can satisfy this definition.
Pf/ Suppose an encryption scheme satisfies the definition.
Consider a distribution $\mathcal{M}$ such that when $m_1 \neq m_2$, $\Pr[M_1 = m_1 \wedge M_2 = m_2] \neq 0$.
Then consider $c_1 = c_2 = c$. By the choice of the distribution $\mathcal{M}$, and the supposition,

$$\Pr[M_1 = m_1 \wedge M_2 = m_2 | C_1 = c \wedge C_2 = c] \neq 0.$$

This means that $\exists$ some k such that $\text{Dec}_k(c) = m_1$ and $\text{Dec}_k(c) = m_2$.
Then consider the first equation, $\text{Dec}_k(c) = m_1$ means that $c = \text{Enc}_k(m_1)$.
Then plug this into the second equation, $\text{Dec}_k(c) = \text{Dec}_k(\text{Enc}_k(m_1)) = m_2$, which breaks the correctness of the scheme.
Thus, no encryption scheme can satisfy this definition.

b

BLANK.

Tuesday

Statistics

Work with Simon

Derive 4.14

Need to show Cov(X,Y) = E[XY]-$\mu_X\mu_Y$ = E[XY]-E[X]E[Y]

$$\begin{aligned}
\text{Cov}(X,Y) &= \text{E}[(X-\mu_X)(Y-\mu_Y)] \\
&= \text{E}[XY + \mu_X\mu_Y - \mu_X Y - \mu_Y X] \\
&= \text{E}[XY] + \text{E}[\mu_X\mu_Y] - \text{E}[\mu_X Y] - \text{E}[\mu_Y X] \\
&= \text{E}[XY] + \mu_X\mu_Y - \mu_X\text{E}[Y] - \mu_Y\text{E}[X] \\
&= \text{E}[XY] + \mu_X\mu_Y - \mu_X\mu_Y - \mu_X\mu_Y \\
&= \text{E}[XY] - \mu_X\mu_Y \\
&= \text{E}[XY] - \text{E}(X)\text{E}(Y)
\end{aligned}$$

X,Y are independent $\Leftrightarrow$ they are uncorrelated.

$\Rightarrow$ Pf/As X, Y are independent, P(XY)=P(X)P(Y).

$$\begin{aligned}
\text{E}[XY] &= \sum_{x,y} xy\text{P}(XY) \\
&= \sum_{x,y} xy\text{P}(X)\text{P}(Y) \\
&= \sum_{x} x\text{P}(X) \sum_{y} y\text{P}(Y) \\
&= E[X]E[Y]
\end{aligned}$$

Therefore by Theorem 4.14, Cov(X,Y) = E[XY]-E[X]E[Y] = 0, which means that X, Y are uncorrelated.

$\Leftarrow$ DisPf/ Counterexample: Let X be uniformly distributed on $\{-1, 0, 1\}$. Let $Y = X^2$. Then,

$$\text{Cov}(X,Y) = \text{Cov}(X,X^2) = \text{E}[X^3] - \text{E}[X]\text{E}[X^2] = 0 - 0 = 0,$$

which means that X and Y are uncorrelated. However, Y is dependent on X. Therefore, this direction is incorrect.

4.27

To find the variance of sum of n independent tetrahedron dice rolls, first find the expectation of a single roll,

$$\begin{aligned} E(X) &= \sum_x x \times P(X = x) \\ &= \frac{1}{4} \times (1 + 2 + 3 + 4) = \frac{5}{2} \end{aligned}$$

As well as that of $E(X^2)$.

$$\begin{aligned} E(X^2) &= \sum_x x^2 \times P(X = x) \\ &= \frac{1}{4} \times (1 + 4 + 9 + 16) = \frac{15}{2} \end{aligned}$$

Then, using these values, find the variance for a single trial.

$$\begin{aligned} V(X) &= E((X - E(X))^2) \\ &= E(X^2) - (E(X)^2) \\ &= \frac{15}{2} - \frac{5}{2}^2 \\ &= \frac{30}{4} - \frac{25}{4} \\ &= \frac{5}{4} \end{aligned}$$

Since the rolls are independent, sum of the variances is the variance of the sum. So for n independent tetrahedron dice rolls, let $Y =$ n independent tetrahedron dice rolls, and $X_i = i^{th}$ tetrahedron dice roll,

$$\begin{aligned} V(Y) &= V(X_1 + X_2 + ...X_n) \\ &= V(X_1) + V(X_2) + ... + V(X_n) \\ &= \frac{5}{4}n \end{aligned}$$

4.37

Since we know X and Y are uncorrelated, $\text{Cov}(X, Y) = 0$. It follows that

$$\begin{aligned}
\text{Var}(Z) &= \text{Var}(wX + (1-w)Y) \\
&= \text{Var}(wX) + \text{Var}((1-w)Y) \\
&= w^2\text{Var}(X) + (1-w)^2\text{Var}(Y) \\
&= w^2\sigma_x^2 + (1-w)^2\sigma_y^2 \\
&= w^2\sigma_x^2 + \sigma_y^2 + w^2\sigma_y^2 - 2w\sigma_y^2 \\
&= (\sigma_x^2 + \sigma_y^2)w^2 - 2w\sigma_y^2 + \sigma_y^2
\end{aligned}$$

Then, $\text{Var}(Z)$ is a function of w and to minimize it, we need to consider the derivative of it with respect to w, which is $\frac{d\text{Var}(Z)}{dw} = 2(\sigma_x^2+\sigma_y^2)w - 2\sigma_y^2$. When the derivative equals to 0, $w = \frac{\sigma_y^2}{\sigma_x^2+\sigma_y^2}$. Evaluating $\frac{d^2\text{Var}(Z)}{dw^2}$ gives that $\text{Var}(Z)$ is concave down, therefore $w = \frac{\sigma_y^2}{\sigma_x^2+\sigma_y^2}$ gives the minimum of $\text{Var}(Z)$.

4.48

X: the first of two fair die rolls.
M: the maximum of the two rolls.
Y: the second of two fair die rolls.

(a)

$$\text{P}(M = m|X = 1) = \begin{cases}
\frac{\text{P}(M=1,X=1)}{\text{P}(X=1)} = \frac{\text{P}(Y=1,X=1)}{\text{P}(X=1)} = \frac{1/6\times 1/6}{1/6} = \frac{1}{6}, & m = 1 \\
\frac{\text{P}(M=2,X=1)}{\text{P}(X=1)} = \frac{\text{P}(Y=2,X=1)}{\text{P}(X=1)} = \frac{1/6\times 1/6}{1/6} = \frac{1}{6}, & m = 2 \\
\frac{1}{6}, & m = 3 \\
\frac{1}{6}, & m = 4 \\
\frac{1}{6}, & m = 5 \\
\frac{1}{6}, & m = 6
\end{cases}$$

$$\text{P}(M = m|X = 2) = \begin{cases}
\frac{\text{P}(M=1,X=2)}{\text{P}(X=2)} = \frac{0}{1/6} = 0, & m = 1 \\
\frac{\text{P}(M=2,X=2)}{\text{P}(X=2)} = \frac{\text{P}(Y=2,X=2)+\text{P}(Y=1,X=2)}{\text{P}(X=2)} = \frac{2/6\times 1/6}{1/6} = \frac{2}{6}, & m = 2 \\
\frac{\text{P}(M=3,X=2)}{\text{P}(X=2)} = \frac{\text{P}(Y=3,X=2)}{\text{P}(X=2)} = \frac{1/6\times 1/6}{1/6} = \frac{1}{6}, & m = 3 \\
\frac{1}{6}, & m = 4 \\
\frac{1}{6}, & m = 5 \\
\frac{1}{6}, & m = 6
\end{cases}$$

$$\mathrm{P}(M=m|X=3)=\begin{cases}\frac{\mathrm{P}(M=1,X=3)}{\mathrm{P}(X=3)}=\frac{0}{1/6}=0, & m=1\\ \frac{\mathrm{P}(M=2,X=3)}{\mathrm{P}(X=3)}=0, & m=2\\ \frac{\mathrm{P}(M=3,X=3)}{\mathrm{P}(X=3)}=\frac{3/6\times 1/6}{1/6}=\frac{3}{6}, & m=3\\ \frac{1}{6}, & m=4\\ \frac{1}{6}, & m=5\\ \frac{1}{6}, & m=6\end{cases}$$

$$\mathrm{P}(M=m|X=4)=\begin{cases}0, & m=1\\ 0, & m=2\\ 0, & m=3\\ \frac{4}{6}, & m=4\\ \frac{1}{6}, & m=5\\ \frac{1}{6}, & m=6\end{cases}$$

$$\mathrm{P}(M=m|X=5)=\begin{cases}0, & m=1\\ 0, & m=2\\ 0, & m=3\\ 0, & m=4\\ \frac{5}{6}, & m=5\\ \frac{1}{6}, & m=6\end{cases}$$

$$\mathrm{P}(M=m|X=6)=\begin{cases}0, & m=1\\ 0, & m=2\\ 0, & m=3\\ 0, & m=4\\ 0, & m=5\\ \frac{6}{6}, & m=6\end{cases}$$

In general,

$$\mathrm{P}(M=m|X=x)=\begin{cases}0, & m<x\\ \frac{x}{6}, & m=x\\ \frac{1}{6}, & m>x\end{cases}$$

(b)

From the equation $\mathrm{E}(M|X=x) = \sum_m m\mathrm{P}(M=m|X=x)$ and cases above, we can get

$$\mathrm{P}(M|X=x) = \begin{cases} \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2}, & x=1 \\ \frac{2}{6}\times 2 + \frac{1}{6}(3+4+5+6) = \frac{22}{6} = \frac{11}{3}, & x=2 \\ \frac{3}{6}\times 3 + \frac{1}{6}(4+5+6) = \frac{24}{6} = 4, & x=3 \\ \frac{4}{6}\times 4 + \frac{1}{6}(5+6) = \frac{27}{6} = \frac{9}{2}, & x=4 \\ \frac{5}{6}\times 5 + \frac{1}{6}(6) = \frac{31}{6}, & x=5 \\ \frac{6}{6}\times 6 = 6, & x=6 \end{cases}$$

. In general,

$$\mathrm{P}(M|X=x) = \frac{21+0+...+(x-1)}{6} = \frac{21+\frac{(x-1)x}{2}}{6}.$$

.

(c)

$$\mathrm{P}(M=m, X=x) = \mathrm{P}(M=m|X=x)\mathrm{P}(X=x) = \begin{cases} 0, & m<x \\ \frac{x}{6}\cdot\frac{1}{6} = \frac{x}{36}, & m=x \\ \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}, & m>x \end{cases}$$

Cryptography

Work with Simon

5. Reed Students' Major

(a)

This is a histogram query. Since each student can only choose one major, students in each major are disjoint, and the addition or removal of a single database element can affect the count in exactly one major, and the difference to that major is bounded by 1. So it has sensitivity 1 and then the database can be released by adding independent draws from $\mathrm{Lap}(1/\epsilon) = \mathrm{Lap}(1/0.7) = \mathrm{Lap}(10/7)$.

(b)

This is acceptable while maintaining privacy because it is post-processing. The additional rounding to zero step is independent with the previous laplace mechanism, and then as the output

released after laplace mechanism is differentially private, the output released after rounding to zero and laplace mechanism is also differentially private.

(c)

The school can release both the number of students per major and the number of students per year by adding noise scaled to $2/\epsilon = 2/0.7$ to the true answer to each query, because the sensitivity in this case is 2 (changing a single database element can affect at most two counts in this situation).

(d)

Using Exponential Mechanism: "What is the largest major?"
Let $\mathcal{X}$ be all majors, drawn from database x. Let $\mathcal{R} = \mathcal{X}$. Let $u(x, r)$, the utility function be the number of people in x with major $r \in \mathcal{R}$. Δu is 1, since we assume that each person can only have one major, and thus contribute to only one count.
Thus, let $M_E(x, u, \mathcal{R})$ be the exponential mechanism that releases the largest major by outputting $r \in \mathcal{R}$ with probability $\propto e^{\epsilon u(x,r)/2\Delta u} = e^{0.7u(x,r)/2}$.

Accuracy: Which one is more accurate, exponential mechanism or laplace mechanism?

(e)

Method 1:
Laplace Mechanism: Consider major i, which has n_i students and each student's year is x_i. Then the expected year of a student in this year is

$$\overline{x_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j$$

$$
\begin{aligned}
\Delta f &= ||\overline{x_i} - \overline{y_i}|| \\
&= ||\frac{1}{n_i}\sum_{j=1}^{n_i} x_j - \frac{1}{n_i - 1}\sum_{j=1}^{n_i - 1} x_j|| \\
&= ||\left(\frac{1}{n_i} - \frac{1}{n_i - 1}\right)\sum_{j=1}^{n_i - 1} x_j + \frac{1}{n_i}x^*|| \\
&\leq ||\left(\frac{-1}{n_i(n_i - 1)}\right) 1 \cdot (n_i - 1) + \frac{4}{n_i}|| \\
&= \frac{3}{n_i}
\end{aligned}
$$

Then for majors with more than 2 students, the sensitivity is bounded by 3 ($n_i = 1$ is taken) after considering neighboring cases. Outputs are released by adding independent draws from Lap($3/\epsilon$). But the information released is not accurate for both big size majors and small size majors

This method needs more information than I currently learnt, like propose-text-release, smooth sensitivity. So... to be continued.

Method 2:
Laplace Mechanism to output both the sum of years of students in one major and the number of students in that major.

The sum of years of students can be released by adding independent draws from Lap($4/\epsilon$) as the sensitivity is 4. The number of students in that major can be released by adding independent draws from Lap($1/\epsilon$) as the sensitivity is 1.

Then the outputs for each major will be a vector (sum of years in the major (noise added), the number of students in that major (noise added)). These outputs are differentially private and will be comparatively accurate when the size of the major is large.

6. Fair statements with $\epsilon = 0.2$

(a) "I promise that no one will learn much about you from the results of my study."

This is not reasonable. For example, if Alice is known to smoke and the result of my study shows that people who smoke will have higher probability of getting cancer, then I know that

Alice has a higher probability of having cancer. In this sense, I know much about Alice from the results of my study.

(b) "I promise that your participation in the study will not cause anyone to learn much about you."

This is reasonable. As my release result is differentially private, your participation in the study should not influence the release result much. Therefore your participation will not provide specific information about you .

(c) "I promise that no one will be able to tell whether you participated in my study."

This is reasonable. My release result is differentially private, which ensures that your participation will not influence the release result much. Thus no one will tell whether you participated or not.

7. 500 smokers with a people smokes

(a) Release a

This is not differentially private.
Consider $\Pr[M(a) = a]$ and $\Pr[M(a+1) = a]$. $\Pr[M(a) = a] = 1$ and $\Pr[M(a+1) = a] = 0$, and therefore there is no ε that can make $1 = \Pr[M(a) = a] \leq e^{\varepsilon}\Pr[M(a+1) = a] = 0$.

(b) Release $a+1$ or $a-1$ with equal probability

This is not differentially private.
Consider $\Pr[M(a) \in \{a+1, a-1\}]$ and $\Pr[M(a+1) \in \{a+1, a-1\}]$. $\Pr[M(a) \in \{a+1, a-1\}]$ $= 1$ and $\Pr[M(a+1) \in \{a+1, a-1\}] = 0$, and therefore there is no ε that can make $1 = \Pr[M(a) \in \{a+1, a-1\}] \leq e^{\varepsilon}\ \Pr[M(a+1) \in \{a+1, a-1\}] = 0$.

(c) Release a or $a-1$ with equal probability

This is not differentially private.
Consider $\Pr[M(a) \in \{a-1\}]$ and $\Pr[M(a+1) \in \{a-1\}]$. $\Pr[M(a) \in \{a-1\}] = 1/2$ and $\Pr[M(a+1) \in \{a-1\}] = 0$, and therefore there is no ε that can make $1/2 = \Pr[M(a) \in \{a-1\}]$ $\leq e^{\varepsilon}\ \Pr[M(a+1) \in \{a-1\}] = 0$.

(d) Release a with prob. $1/2$ or $a+1$ or $a-1$ with prob. $1/4$

This is not differentially private.
Consider $\Pr[M(a) \in \{a-1\}]$ and $\Pr[M(a+1) \in \{a-1\}]$. $\Pr[M(a) \in \{a-1\}] = 1/4$ and $\Pr[M(a+1) \in \{a-1\}] = 0$, and therefore there is no ε that can make $1/4 = \Pr[M(a) \in \{a-1\}] \leq e^{\varepsilon} \Pr[M(a+1) \in \{a-1\}] = 0$.

(e) For each person, flip his answer with prob. p

i. $p = 0$

This is not differentially private. It is the same as part(a) to output a directly.

ii. $p = 0.1$

iii. $p = 0.5$

iv. $p = 1$

This is not differentially private.
Consider $\Pr[M(a) = 500 - a]$ and $\Pr[M(a+1) = 500 - a]$. $\Pr[M(a) = 500 - a] = 1$ and $\Pr[M(a+1) = 500 - a] = 0$, and therefore there is no ε that can make $1 = \Pr[M(a) = 500 - a] \leq e^{\varepsilon} \Pr[M(a+1) = 500 - a] = 0$.

Wednesday

Statistics

1. Sum of two independent exponential random variables

Step 1: As both X and Y are exponential random variables, $0 \leq X \leq \infty$ and $0 \leq Y \leq \infty$. Therefore $0 \leq W = X + Y \leq \infty$.

Step 2:

$$\begin{aligned}
\mathrm{F}_W(w) &= \mathrm{P}(W \le w) = \mathrm{P}(X + Y \le w) = \mathrm{P}(X \le w - Y) \\
&= \int_0^w \int_0^{w-y} f(x,y)\ dx\ dy \\
&= \int_0^w \int_0^{w-y} f(x)f(y)\ dx\ dy \\
&= \int_0^w \int_0^{w-y} \lambda_X e^{-\lambda_X x} \lambda_Y e^{-\lambda_Y y}\ dx\ dy \\
&= \frac{(\lambda_X - \lambda_Y) + (\lambda_Y e^{-w\lambda_X} - \lambda_X e^{-w\lambda_Y})}{\lambda_X - \lambda_Y}
\end{aligned}$$

Step 3:

$$\begin{aligned}
\frac{\partial \mathrm{F}_W(w)}{\partial w} &= \partial \left(\frac{(\lambda_X - \lambda_Y) + (\lambda_Y e^{-w\lambda_X} - \lambda_X e^{-w\lambda_Y})}{\lambda_X - \lambda_Y} \right) / \partial w \\
&= \frac{\lambda_X \lambda_Y (e^{-w\lambda_Y} - e^{-w\lambda_X})}{\lambda_X - \lambda_Y} \\
&= f_W(w)
\end{aligned}$$

Step 4: If $\lambda_X = \lambda_Y = \lambda$, the above equation does not make sense. So we will reevaluate this case. From step 2,

$$\begin{aligned}
\mathrm{F}_W(w) &= \int_0^w \int_0^{w-y} \lambda_X e^{-\lambda_X x} \lambda_Y e^{-\lambda_Y y}\ dx\ dy \\
&= \int_0^w \int_0^{w-y} \lambda e^{-\lambda x} \lambda e^{-\lambda y}\ dx\ dy \\
&= 1 - e^{-w\lambda}(1 + w\lambda)
\end{aligned}$$

Then do the same thing in step 3,

$$\begin{aligned}
\frac{\partial \mathrm{F}_W(w)}{\partial w} &= \partial(1 - e^{-w\lambda}(1 + w\lambda))/\partial w \\
&= \lambda^2 w e^{-\lambda w} \\
&= f_W(w)
\end{aligned}$$

Step 5: Combing result in step 3 and step 4,

$$f_W(w) = \begin{cases} \lambda^2 w e^{-\lambda w}, & \lambda_X = \lambda_Y = \lambda \\ \frac{\lambda_X \lambda_Y (e^{-w\lambda_Y} - e^{-w\lambda_X})}{\lambda_X - \lambda_Y}, & \lambda_X \neq \lambda_Y \end{cases}$$

2. Square of normal random variable $Z \sim \mathbf{Norm}(0, 1)$

Step 1:

$$\begin{aligned}
\mathrm{F}_X(x) &= \mathrm{P}(X \le x) = \mathrm{P}(Z^2 \le x) \\
&= \mathrm{P}(-\sqrt{x} \le Z \le \sqrt{x}) \\
&= \mathrm{P}(Z \le \sqrt{x}) - \mathrm{P}(Z \le -\sqrt{x}) \\
&= \mathrm{F}_Z(\sqrt{x}) - \mathrm{F}_Z(-\sqrt{x})
\end{aligned}$$

Step 2:

$$\begin{aligned}
\frac{\partial}{\partial x}\mathrm{F}_X(x) &= \frac{\partial}{\partial x}\left(\mathrm{F}_Z(\sqrt{x}) - \mathrm{F}_Z(-\sqrt{x})\right) \\
&= f_Z(\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}} - (-f_Z(-\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}}) \\
&= f_Z(\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}} + f_Z(\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}} \\
&= f_Z(\sqrt{x}) \cdot \frac{1}{\sqrt{x}} \\
&= \frac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}x} \\
&= f_X(x)
\end{aligned}$$

In conclusion, $f_X(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}x}$.

Thursday

Statistics

Problem 1A

Given $X \sim \mathrm{Norm}(\mu, \sigma^2)$, and $\bar{X}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_{ij}$.

We know that X_{ij} is independent on each other, and we know that $m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$ for all X_{ij}, which is the moment generating function of normal distribution.

$$
\begin{aligned}
m_{\bar{X}_j}(t) &= m_{\frac{1}{n_j}\sum_{i=1}^{n_j} X_{ij}}(t) \\
&= m_{\sum_{i=1}^{n_j} X_{ij}}(\frac{t}{n_j}) \\
&= m_{X_{1j}}(\frac{t}{n_j}) m_{X_{2j}}(\frac{t}{n_j}) ... m_{X_{n_j j}}(\frac{t}{n_j}) \\
&= (e^{\mu(t/n_j)+\frac{\sigma^2(t/n_j)^2}{2}})^{n_j} \\
&= e^{\mu t+\frac{\sigma^2 t^2}{2n_j}}
\end{aligned}
$$

Therefore, $\bar{X}_j \sim \text{Norm}(\mu, \frac{\sigma^2}{n_j})$.

Problem 1B

First, We know the following things:
(a) From the condition given, $X_{ij} \sim \text{Norm}(\mu, \sigma^2)$. After standalization leads to $\frac{X_{ij}-\mu}{\sigma} \sim \text{Norm}(\mu, \sigma^2)$.
(b) From Problem 1A, $\bar{X}_j \sim \text{Norm}(\mu, \frac{\sigma^2}{n_j})$. After standalization leads to $(\frac{\bar{X}_j-\mu}{\sigma})\sqrt{n_j} \sim \text{Norm}(0,1)$.
(c) From Wednesday Problem 2, if $Z \sim \text{Norm}(0,1)$, $Z^2 \sim \mathcal{X}^2(1)$.
(d) The moment generating function of the sum of independent variables is the product of the moment generating function of each independent variable.
(e) Lemma 1 and Lemma 2 about the independence between sample mean and sample variance.

Then consider $\sum_{i=1}^{n_j}(X_{ij} - \bar{X}_j)^2$.

$$\begin{aligned}
\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2 &= \sum_{i=1}^{n_j}(X_{ij}-\mu+\mu-\bar{X}_j)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 + 2(\mu-\bar{X}_j)\sum_{i=1}^{n_j}(X_{ij}-\mu) + n_j(\mu-\bar{X}_j)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 + 2(\mu-\bar{X}_j)\left(-(n_j\mu)+\sum_{i=1}^{n_j}X_{ij}\right) + n_j(\bar{X}_j-\mu)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 + 2(\mu-\bar{X}_j)\left(-(n_j\mu)+n_j\bar{X}_j\right) + n_j(\bar{X}_j-\mu)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 - 2n_j(\mu-\bar{X}_j)(\mu-\bar{X}_j) + n_j(\bar{X}_j-\mu)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 - 2n_j(\mu-\bar{X}_j)^2 + n_j(\bar{X}_j-\mu)^2 \\
&= \sum_{i=1}^{n_j}(X_{ij}-\mu)^2 - n_j(\mu-\bar{X}_j)^2 \\
&= \sigma^2\left(\sum_{i=1}^{n_j}\frac{(X_{ij}-\mu)^2}{\sigma^2} - n_j\frac{(\mu-\bar{X}_j)^2}{\sigma^2}\right)
\end{aligned}$$

It follows that (I need this one because I do not know how to show that $\sum_{i=1}^{n_j}\frac{(X_{ij}-\mu)^2}{\sigma^2}$ is independent of $n_j\frac{(\mu-\bar{X}_j)^2}{\sigma^2}$. I only can get $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2$ is independent of $n_j\frac{(\mu-\bar{X}_j)^2}{\sigma^2}$ which will be talked about later.)

$$\sum_{i=1}^{n_j}\frac{(X_{ij}-\mu)^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2 + n_j\frac{(\mu-\bar{X}_j)^2}{\sigma^2}.$$

First consider $\sum_{i=1}^{n_j}\frac{(X_{ij}-\mu)^2}{\sigma^2}$. As X_{ij} are independent of each other, no matter which set j we are looking at, X_{ij} should be independent variables of each other. It follows that $\frac{(X_{ij}-\mu)^2}{\sigma^2}$ are independent variables for different i as μ and σ are just constants.

Also by (a), we know that $\frac{X_{ij}-\mu}{\sigma}\sim \text{Norm}(\mu,\sigma^2)$. Then by (c), we can get $\frac{(X_{ij}-\mu)^2}{\sigma^2}\sim\mathcal{X}^2(1)$ for all i. As the moment generating function for gamma distribution is $(1-\frac{t}{\beta})^{-\alpha}$, by substituting specific value of α and β, the moment generating function for $\mathcal{X}^2(1)$ is $(1-2t)^{-\frac{1}{2}}$. It follows

that,

$$\begin{aligned} m_{\sum_{i=1}^{n_j} \frac{(X_{ij}-\mu)^2}{\sigma^2}}(t) &= \left(m_{\frac{(X_{1j}-\mu)^2}{\sigma^2}}(t)\right)\left(m_{\frac{(X_{2j}-\mu)^2}{\sigma^2}}(t)\right)...\left(m_{\frac{(X_{n_jj}-\mu)^2}{\sigma^2}}(t)\right) \\ &= ((1-2t)^{-\frac{1}{2}})^{n_j} \\ &= (1-2t)^{-\frac{n_j}{2}} \end{aligned}$$

Then consider $n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}$. By (b) and (c), we can get $n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2} = (\sqrt{n_j}\frac{(\mu-\bar{X_j})}{\sigma})^2 \sim \mathcal{X}^2(1)$. Thus, $m_{n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}}(t) = (1-2t)^{-\frac{1}{2}}$.

By (e), $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2$ is independent of $n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}$. Therefore

$$\begin{aligned} m_{\sum_{i=1}^{n_j} \frac{(X_{ij}-\mu)^2}{\sigma^2}}(t) &= m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 + n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}}(t) \\ &= m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2}(t) \cdot m_{n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}}(t) \end{aligned}$$

It follows that

$$\begin{aligned} m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2}(t) &= \frac{m_{\sum_{i=1}^{n_j} \frac{(X_{ij}-\mu)^2}{\sigma^2}}(t)}{m_{n_j\frac{(\mu-\bar{X_j})^2}{\sigma^2}}(t)} \\ &= \frac{(1-2t)^{-\frac{n_j}{2}}}{(1-2t)^{-\frac{1}{2}}} \\ &= (1-2t)^{-\frac{n_j-1}{2}} \end{aligned}$$

Thus, $\frac{1}{\sigma^2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 \sim \mathcal{X}^2(n_j-1)$, which means that $\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 \sim \sigma^2\mathcal{X}^2(n_j-1)$.

Then consider $\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 = \sigma^2\frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2$.
As all i and j, X_{ij} are independent of each other, and for all j, $\bar{X_j}$ are independent, then for all i and j, $\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2$ are independent, then we can apply rule (d).

$$\begin{aligned} m_{\frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2}(t) &= m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_{i1}-\bar{X_1})^2}(t) \cdot m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_2}(X_{i2}-\bar{X_2})^2}(t)...m_{\frac{1}{\sigma^2}\sum_{i=1}^{n_k}(X_{ik}-\bar{X_k})^2}(t) \\ &= (1-2t)^{-\frac{n_1-1}{2}} \cdot (1-2t)^{-\frac{n_2-1}{2}}... \cdot (1-2t)^{-\frac{n_k-1}{2}} \\ &= (1-2t)^{-\left(\frac{n_1-1}{2}+\frac{n_2-1}{2}+..+\frac{n_k-1}{2}\right)} \\ &= (1-2t)^{-\frac{n_1+...+n_k-k}{2}} \\ &= (1-2t)^{-\frac{N-k}{2}} \end{aligned}$$

Thus, $\frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 \sim \mathcal{X}^2(N-k)$, which means that $W = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X_j})^2 \sim \sigma^2\mathcal{X}^2(N-k)$.

Problem 1C

First consider the distribution of $\bar{X} = \frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij}$ (Note that X_{ij} are independent of each other for all i and j).
Calculation similar to 1A gives

$$\begin{aligned} m_{\sum_{i=1}^{n_j} X_{ij}}(t) &= m_{X_{1j}}(t) m_{X_{2j}}(t) ... m_{X_{n_j j}}(t) \\ &= (e^{\mu t + \frac{\sigma^2 t^2}{2}})^{n_j} \\ &= e^{\mu t n_j + \frac{\sigma^2 t^2 n_j}{2}} \end{aligned}$$

It follows that

$$\begin{aligned} m_{\bar{X}}(t) &= m_{\frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij}}(t) \\ &= m_{\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij}}(\frac{t}{N}) \\ &= m_{\sum_{i=1}^{n_1} X_{i1}}(\frac{t}{N}) \cdot m_{\sum_{i=1}^{n_2} X_{i2}}(\frac{t}{N}) \cdot ... \cdot m_{\sum_{i=1}^{n_j} X_{in_j}}(\frac{t}{N}) \\ &= e^{\mu \frac{t}{N} n_1 + \frac{\sigma^2 \left(\frac{t}{N}\right)^2 n_1}{2}} \cdot e^{\mu \frac{t}{N} n_2 + \frac{\sigma^2 \left(\frac{t}{N}\right)^2 n_2}{2}} \cdot ... \cdot e^{\mu \frac{t}{N} n_k + \frac{\sigma^2 \left(\frac{t}{N}\right)^2 n_k}{2}} \\ &= e^{\mu \frac{t}{N}(n_1+n_2+...+n_k) + \frac{\sigma^2 \left(\frac{t}{N}\right)^2 (n_1+n_2+...+n_k)}{2}} \\ &= e^{\mu \frac{t}{N} N + \frac{\sigma^2 \left(\frac{t}{N}\right)^2 N}{2}} \\ &= e^{\mu t + \frac{\sigma^2 t^2}{2N}} \end{aligned}$$

Thus, $\bar{X} \sim \text{Norm}(\mu, \frac{\sigma^2}{N})$. After standalization, we get $\sqrt{N}\frac{\bar{X}-\mu}{\sigma} \sim \text{Norm}(0,1)$. Then from 1B rule (c), we get $N\frac{(\bar{X}-\mu)^2}{\sigma^2} \sim \mathcal{X}^2(1)$.

Then consider $B = \sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X})^2$.

$$
\begin{aligned}
B &= \sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X})^2 \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu + \mu - \bar{X})^2 \\
&= \sum_{j=1}^{k} n_j\left((\bar{X}_j - \mu)^2 + 2(\bar{X}_j - \mu)(\mu - \bar{X}) + (\mu - \bar{X})^2\right) \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 + 2\sum_{j=1}^{k} n_j(\bar{X}_j - \mu)(\mu - \bar{X}) + \sum_{j=1}^{k} n_j(\mu - \bar{X})^2 \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - 2(\bar{X} - \mu)\sum_{j=1}^{k} n_j(\bar{X}_j - \mu) + (\mu - \bar{X})^2\sum_{j=1}^{k} n_j \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - 2(\bar{X} - \mu)\left(\left(\sum_{j=1}^{k} n_j\bar{X}_j\right) - \mu\sum_{j=1}^{k} n_j\right) + (\mu - \bar{X})^2\sum_{j=1}^{k} n_j \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - 2(\bar{X} - \mu)\left(\left(\sum_{j=1}^{k} n_j\bar{X}_j\right) - \mu\sum_{j=1}^{k} n_j\right) + (\mu - \bar{X})^2\sum_{j=1}^{k} n_j
\end{aligned}
$$

As $\bar{X} = \frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij} = \frac{1}{N}\sum_{j=1}^{k} n_j\bar{X}_j$, $\sum_{j=1}^{k} n_j\bar{X}_j = N\bar{X}$.

Then we can keep working on the above equation,

$$
\begin{aligned}
B &= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - 2(\bar{X} - \mu)\left(N\bar{X} - \mu N\right) + (\mu - \bar{X})^2 N \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - 2N(\bar{X} - \mu)^2 + (\mu - \bar{X})^2 N \\
&= \sum_{j=1}^{k} n_j(\bar{X}_j - \mu)^2 - N(\bar{X} - \mu)^2
\end{aligned}
$$

Thus, $\frac{1}{\sigma^2}B = \sum_{j=1}^{k} n_j\frac{(\bar{X}_j - \mu)^2}{\sigma^2} - N\frac{(\bar{X} - \mu)^2}{\sigma^2}$, which is

$$
\sum_{j=1}^{k} n_j\frac{(\bar{X}_j - \bar{X})^2}{\sigma^2} = \sum_{j=1}^{k} n_j\frac{(\bar{X}_j - \mu)^2}{\sigma^2} - N\frac{(\bar{X} - \mu)^2}{\sigma^2}
$$

.

From 1B, $m_{n_j \frac{(\mu-\bar{X}_j)^2}{\sigma^2}}(t) = (1-2t)^{-\frac{1}{2}}$. It follows that

$$\begin{aligned} m_{\sum_{j=1}^{k} n_j \frac{(\bar{X}_j-\mu)^2}{\sigma^2}}(t) &= m_{n_1 \frac{(\bar{X}_1-\mu)^2}{\sigma^2}}(t) \cdot m_{n_2 \frac{(\bar{X}_2-\mu)^2}{\sigma^2}}(t) \cdot \ldots \cdot m_{n_k \frac{(\bar{X}_k-\mu)^2}{\sigma^2}}(t) \\ &= \left((1-2t)^{-\frac{1}{2}}\right)^k \\ &= (1-2t)^{-\frac{k}{2}} \end{aligned}$$

As talked above that $N\frac{(\bar{X}-\mu)^2}{\sigma^2} \sim \mathcal{X}^2(1)$, $m_{N\frac{(\bar{X}-\mu)^2}{\sigma^2}} = (1-2t)^{-\frac{1}{2}}$.

Then by the same argument as in 1B, $\sum_{j=1}^{k} n_j \frac{(\bar{X}_j-\bar{X})^2}{\sigma^2}$ is independent of $N\frac{(\bar{X}-\mu)^2}{\sigma^2}$. Also, by same process of change in formula in 1B, we can get

$$\begin{aligned} m_{\frac{1}{\sigma^2}B}(t) &= m_{\sum_{j=1}^{k} n_j \frac{(\bar{X}_j-\bar{X})^2}{\sigma^2}}(t) \\ &= \frac{m_{\sum_{j=1}^{k} n_j \frac{(\bar{X}_j-\mu)^2}{\sigma^2}}(t)}{m_{N\frac{(\bar{X}-\mu)^2}{\sigma^2}}(t)} \\ &= \frac{(1-2t)^{-\frac{k}{2}}}{(1-2t)^{-\frac{1}{2}}} \\ &= (1-2t)^{-\frac{k-1}{2}} \end{aligned}$$

Therefore $\frac{1}{\sigma^2}B \sim \mathcal{X}^2(k-1)$, which is $B \sim \sigma^2\mathcal{X}^2(k-1)$.