HW0-a

Probability

A.1

Denote the event of having disease as D and the event of not having disease as H.
Denote the event of test psitive as + and thevent of test negative as -.
From the question statement, we know that $P(+|D) = 0.99$ then $P(-|D) = 0.01$
Plus, $P(-|H) = 0.99$ then $P(+|H) = 0.01$
Also, $P(D) = 10^{-4}$ then $P(H) = 0.9999$

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|H)P(H)} = \frac{0.99 * 10^{-4}}{0.99 * 10^{-4} + 0.01 * 0.9999} \approx 0.009804$$

A.2

a)

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

To solve the equation above, we want to derive information from $E(Y|X = x)$
According to iterated expectation,
$E(XY) = EE(XY|X) = E[XE(Y|X)] = E[X * X] = E(X^2)$
and $E(Y) = EE(Y|X) = E(X)$

$$Cov(X, Y) = E(X^2) - E(X)E(X) = E[(X - EX)^2]$$

b)

As X and Y are independent, $E(XY) = E(X)E(Y)$
If we derive $Cov(X, Y)$ from the given formula,

$$\begin{aligned} Cov(X, Y) = E[(X - EX)(Y - EY)] \\ = E(XY - EX * Y - XEY + EXEY) \\ = E(XY) - EX * EY - EX * EY + EX * EY \\ = EX * EY - EX * EY - EX * EY + EX * EY = 0 \end{aligned}$$

Or simply

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(X)(Y) = E(X)E(Y) = 0$$

A.3

a)

$$h(z) = h(x + y)$$

Define $H(z)$ as cdf of $h(z)$ then

$$\begin{aligned} H(z) = P(x + y < z) \\ = \int\int_{x+y \leq z} h_{X,Y}(x, y)dxdy \\ = \int_{-\infty}^{\infty} (\int_{\infty}^{z-x} h_{X,Y}(x, y)dy)dx \\ = \int_{\infty}^{\infty} \int_{-\infty}^{z} h_{X,Y}(x, w - x)dwdx \\ = \int_{\infty}^{z} \int_{-\infty}^{\infty} h_{X,Y}(x, w - x)dxdw \end{aligned}$$

According the cdf formula

$$F_{X+Y}(z) = \int_{-\infty}^{z} f_{X+Y}(w)dw$$

The inner integral is actually h_{X+Y}

$$h_{X+Y}(z) = \int_{-\infty}^{\infty} h_{X,Y}(x, z-x)dx$$

As X, Y are independent

$$h_z = \int_{-\infty}^{\infty} f(x)g(z-x)dx$$

b)

As sample space for X,Y are both [0,1], we don't need consider range outside of it.
Therefore,
when $0 < z \leq 1$, $h(z) = \int_0^z 1 \times 1 = z$
when $1 < z < 2$, $h(z) = \int_{z-1}^1 1 \times 1 = 2 - z$
Otherwise it's 0.

A.4

According to the linearity of expectation,

$$E(Y) = E(aX + b) = aE(X) + b$$

Also,

$$Var(Y) = Var(aX + b) = a^2 Var(X) + b$$

According the formula above and the question statement, we can get

$$a\mu + b = 0$$
$$a^2\sigma^2 = 1$$

By soving it, we can get two solutions

solution1:

$$a = \frac{1}{\sigma} \quad b = -\frac{\mu}{\sigma}$$

solution2:

$$a = -\frac{1}{\sigma} \quad b = \frac{\mu}{\sigma}$$

A.5

Expectation

$$E(\sqrt{n}(\hat{\mu}_n - \mu) = \sqrt{n}[E(\hat{\mu}_n) - u]$$

Solve $E(\hat{\mu}_n)$ first

$$E(\hat{\mu}_n) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n} * n * E(X) = E(X)$$

Then we can get

$$E(\sqrt{n}(\hat{\mu}_n - \mu) = \sqrt{n}(EX - EX) = 0$$

Variance

$$Var(\sqrt{n}(\hat{\mu}_n - \mu) = n * Var(\hat{\mu}_n - \mu)$$
$$= n * Var(\frac{1}{n}(X_1 + X_2 + X_3 + ... + X_n)$$

$$= n * \frac{1}{n^2} Var(X_1 + X_2 + X_3 + ... + X_n)$$
$$= \frac{1}{n}(Var(X_1) + Var(X_2) + Var(X_3) + ... + Var(X_n))$$
$$= \frac{1}{n} * n * \sigma^2 = \sigma^2$$

A.6

a)

$$\begin{gathered} E[\hat{F}_n(x)] = E(\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{y \le x\}) \\ = \frac{1}{n}\sum_{i=1}^{n} E(\mathbb{I}\{y \le x\}) \\ = \frac{1}{n}\sum_{i=1}^{n} P(\mathbb{I}\{y \le x\}) \\ = P(\mathbb{I}\{y \le x\}) \\ = E(\mathbb{I}\{y \le x\} \\ = F(x) \end{gathered}$$

b)

$$\begin{gathered} Varaince(\hat{F}_n(x)) = E((\hat{F}_n(x) - F(x))^2) \\ = E(\hat{F}_n(x)^2) - E(\hat{F}_n(x))^2 \end{gathered}$$

Therefore, let's solve $E(\hat{F}_n(x)^2)$ as $E(\hat{F}_n(x))$ is known from a)

$$\begin{gathered} E(\hat{F}_n(x)^2) = E[\frac{1}{n}\sum_{i=1}^{n}(\mathbb{I}\{x_i \le x\})^2] \\ = \frac{1}{n^2}E(\frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\{x_i \le x\}\sum_{j=1}^{n}\mathbb{I}\{x_j \le x\} \\ = \frac{1}{n^2}E(\sum_{i=j=1}^{n}(\mathbb{I}\{x_i \le x\})^2 + \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}(\mathbb{I}\{x_i \le x\})\mathbb{I}\{x_j \le x\})) \\ \text{After pushing E inside} \\ = \frac{1}{n^2}\sum_{i=j=1}^{n}E(\mathbb{I}\{x_i \le x\}) + \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}E(\mathbb{I}\{x_i \le x\})E(\mathbb{I}\{x_j \le x\}) \\ = \frac{1}{n^2}\sum_{i=j=1}^{n}E(\mathbb{I}\{x_i \le x\}) + \sum_{i=1}^{n}E(\mathbb{I}\{x_i \le x\})\sum_{j=1, j\neq i}^{n}E(\mathbb{I}\{x_j \le x\}) \\ = \frac{1}{n^2}n * F(x) + (n^2 - n)F(x)^2 \\ = \frac{F(x) + (n-1)F(x)^2}{n} \end{gathered}$$

Plug the result above and $E(\hat{F}_n(x))$ from a) into $Varaince(\hat{F}_n(x))$, we get

$$Varaince(\hat{F}_n(x)) = \frac{F(x) + (n-1)F(x)^2}{n} - F(x)^2 = \frac{F(x)(1 - F(x))}{n}$$

c)

From b) we know that

$$Varaince(\hat{F}_n(x)) = \frac{F(x)(1 - F(x))}{n}$$

It reaches maximum when $F(x) = 1 - F(x)$ where $F(x)$ is 0.5.
Therefore,

$$Varaince(\hat{F}_n(x)) = \frac{0.5 * (1 - 0.5)}{n} = \frac{1}{4n}$$

B.1

Linear Algebra and Vector Calculus

A.7

a)

First we reduce both matrix

$$A=\begin{bmatrix}1&2&1\\1&0&3\\1&1&2\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1\\0&-2&2\\0&-1&1\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1\\0&-2&2\\0&0&0\end{bmatrix}$$
$$\begin{bmatrix}1&2&3\\1&0&1\\1&1&2\end{bmatrix}\rightarrow\begin{bmatrix}1&2&3\\0&-2&-2\\0&-1&-1\end{bmatrix}\rightarrow\begin{bmatrix}9&2&3\\0&-2&-2\\0&0&0\end{bmatrix}$$

Both matrix has two non-zero leading pivots so their rank are both 2.

b)

basis for A's column span

$$\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}2\\0\\1\end{pmatrix}\}$$

basis for B's column span

$$\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}2\\0\\1\end{pmatrix}\}$$

A.8

a)

$$\begin{bmatrix}0&2&4\\2&4&2\\3&3&1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}0\times1+2\times1+4\times1\\2\times1+4\times1+2\times1\\3\times1+3\times1+1\times1\end{bmatrix}=\begin{bmatrix}6\\8\\7\end{bmatrix}$$

b)

Rewrite $Ax=b$ as $x=A^{-1}b$
Write the equation in augmented matrix

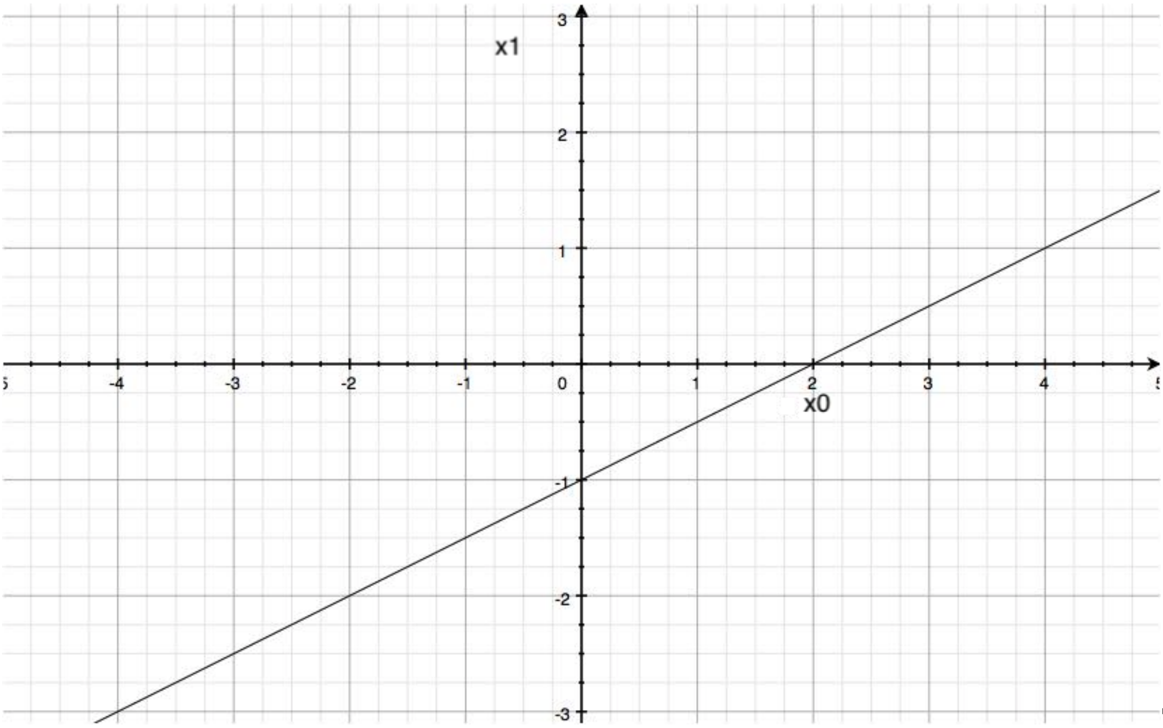
$$\begin{bmatrix}0&2&4&1&0&0\\2&4&2&0&1&0\\3&3&1&0&0&1\end{bmatrix}\rightarrow\begin{bmatrix}2&4&2&0&1&0\\0&2&4&1&0&0\\3&3&1&0&0&1\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1&0&\frac{1}{2}&0\\0&2&4&1&0&0\\3&3&1&0&0&1\end{bmatrix}$$
$$\rightarrow\begin{bmatrix}1&2&1&0&\frac{1}{2}&0\\0&1&2&\frac{1}{2}&0&0\\0&-3&-2&0&\frac{-3}{2}&1\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1&0&\frac{1}{2}0&\\0&1&2&\frac{1}{2}&0&0\\0&0&4&\frac{3}{2}&\frac{-3}{2}&1\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1&0&\frac{1}{2}&0\\0&1&2&\frac{1}{2}&0&0\\0&0&1&\frac{3}{8}&\frac{-3}{8}&\frac{1}{4}\end{bmatrix}$$
$$\rightarrow\begin{bmatrix}1&2&0&\frac{-3}{8}&\frac{7}{8}&\frac{-1}{4}\\0&1&0&\frac{-1}{4}&\frac{3}{4}&\frac{-1}{2}\\0&0&1&\frac{3}{8}&\frac{-3}{8}&\frac{1}{4}\end{bmatrix}\rightarrow\begin{bmatrix}1&2&0&\frac{-3}{8}&\frac{7}{8}&\frac{-1}{4}\\0&1&0&\frac{-1}{4}&\frac{3}{4}&\frac{-1}{2}\\0&0&1&\frac{3}{8}&\frac{-3}{8}&\frac{1}{4}\end{bmatrix}\begin{bmatrix}1&0&0&\frac{1}{8}&\frac{-5}{8}&\frac{3}{4}\\0&1&0&\frac{-1}{4}&\frac{3}{4}&\frac{-1}{2}\\0&0&1&\frac{3}{8}&\frac{-3}{8}&\frac{1}{4}\end{bmatrix}$$

Therefore

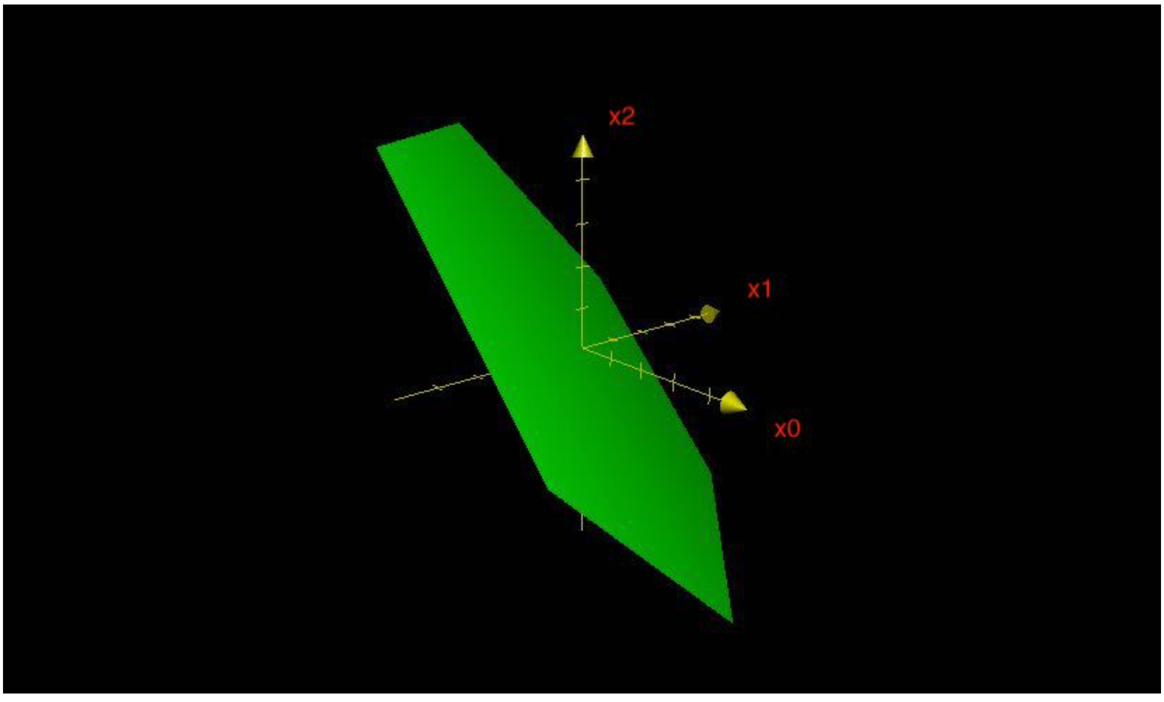
$$x=A^{-1}b=\begin{bmatrix}\frac{1}{8}&\frac{-5}{8}&\frac{3}{4}\\\frac{-1}{4}&\frac{3}{4}&\frac{-1}{2}\\\frac{3}{8}&\frac{-3}{8}&\frac{1}{4}\end{bmatrix}\begin{bmatrix}-2\\-2\\-4\end{bmatrix}=\begin{bmatrix}-2\\1\\-1\end{bmatrix}$$

A.9

a)



b)



c)

From the constraint, we can get $w^T x = -b$.

$$||x_0 - x||^2 = (\frac{w^T x_0 - w^t x}{||w||})^2 = (\frac{w^t x_0 + b}{||w||})^2$$

A.10 (check back later)

a)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} A_{i,j} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n} B_{i,j} y_i x_j + c$$

b)

For $\forall k \in \{1, 2, ..., n\}$

$$\frac{\partial f(x,y)}{\partial x_k} = 2A_{kk}x_k + \sum_{i=1, i\neq k}^{n} A_{i,1}x_i + \sum_{j=1, j\neq k}^{n} A_{1,j}x_j + \sum_{i=1}^{n} B_{i,1} * y_i$$

$$\nabla_x f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x_1} & \frac{\partial f(x,y)}{\partial x_2} & \dots & \frac{\partial f(x,y)}{\partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} 2A_{1,1}x_1 + \sum_{i=2}^{n} A_{i,1}x_i + \sum_{j=2}^{n} A_{1,j}x_j + \sum_{i=1}^{n} B_{i,1} * y_i & \dots & 2A_{n,n} * x_n + \sum_{i=1}^{n-1} A_{i,1}x_i + \sum_{j=1}^{n-1} A_{1,j}x_j + \sum_{i=1}^{n} B_{i,n} * y_i \end{bmatrix}$$

c)

For $\forall k \in \{1, 2, ..., n\}$

$$\frac{\partial f(x,y)}{\partial y_k} = \sum_{j=1}^{n} B_{k,j}x_j$$

$$\nabla_y f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial y_1} & \frac{\partial f(x,y)}{\partial y_2} & \dots & \frac{\partial f(x,y)}{\partial y_n} \end{bmatrix}$$
$$= \begin{bmatrix} \sum_{j=1}^{n} B_{1,j}x_j & \sum_{j=1}^{n} B_{2,j}x_j & \dots & \sum_{j=1}^{n} B_{n,j}x_j \end{bmatrix}$$

B.2

B.3

Programming

A.11

a)

A^{-1}

```
[[ 0.125 -0.625  0.75 ]
 [-0.25   0.75  -0.5  ]
 [ 0.375 -0.375  0.25 ]]
```

b)

$A^{-1}b$

```
[[-2.]
 [ 1.]
 [-1.]]
```

Ac

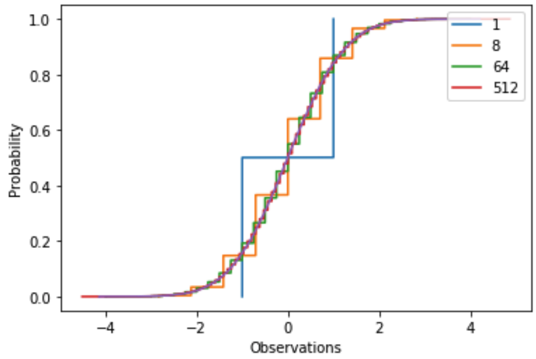
```
[[6]
 [8]
 [7]]
```

Code

```
import numpy as np

A = np.array([[0,2,4],[2,4,2],[3,3,1]])
print("A^(-1)")
print(np.linalg.inv(A))

b = np.array([[-2],[-2],[-4]])
```

```
print("A^(-1)b")
result = np.matmul(np.linalg.inv(A),b)
print(result)

c = np.array([[1],[1],[1]])
print("Ac")
print(np.matmul(A, c) )
```

A.12




```
import math
import matplotlib.pyplot as plt
import numpy as np
ks = [1,8,64,512]
n = round(1/(0.0025**2 *4))
for k in ks:
    ns = np.sum(np.sign(np.random.randn(n, k))*np.sqrt(1./k), axis=1)
    plt.plot(np.sort(ns), np.arange(1, n+1)/float(n), label = str(k))
plt.step(sorted(np.random.randn(n)), np.arange(1, n+1)/float(n))
plt.legend(loc = 1)
plt.xlabel("Observations")
plt.ylabel("Probability")
print(n) # 40000
```