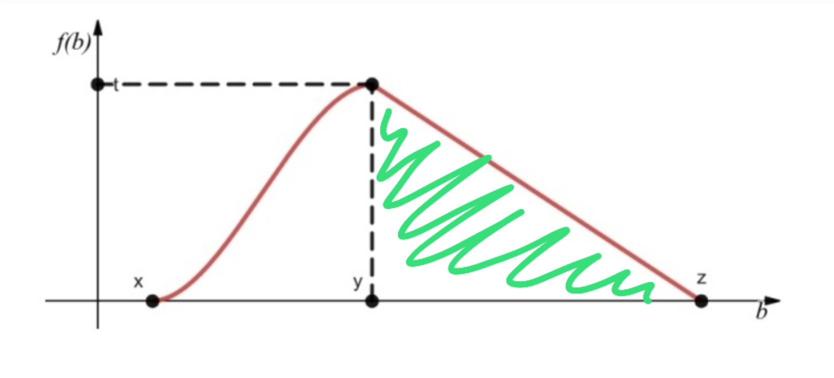
Rana Elif Albayrak
280201037

Task 1

2. Looking at the pmf of A, we can say it looks like geometric distribution.
For a geometric distribution, the probability is calculated as $P(X=x)=(1-p)^{x-1}\cdot p$ where p is the probability of succes in a single trial. In this case, $p=0.5$ for A.
For geometric distribution, X is number of trials and it has infinitely many possible values $\{1,2,3,\dots\}$. But, A is exercise frequency and can take infinitely many possible values $\{0,1,2,\dots\}$. The difference is that, 0 is included for A, but not for X. So, we can define $A=X-1$ in terms of X.

4. $f(b)=-0.096b^3+0.432b^2-0352b+0.08$, $x\le b\le y$
$f(b)=(-2b+11)/15$, $y\le b\le z$
$f(b)=0$, elsewhere



Since $f(b)$ is a linear function between $y\le b\le z$, we can calculate green shaded area as a triangle, which is $\frac{(z-y)\cdot t}{2}$.

By looking at the graph, we can say $f(z)=f(x)=0$.
So, we need to solve $f(z)=\frac{-2z+11}{15}=0$. From this, we get $\boxed{z=\frac{11}{2}=5.5}$

For x, we need to solve $f(x) = -0.096x^3 + 0.432x^2 - 0.352x + 0.08 = 0$

Multiplying both sides by 1000

$-96x^3 + 432x^2 - 352x + 80 = 0$

Organizing common factors

$-16(2x-1)(3x^2-12x+5) = 0$

Zero factor principle: If $ab=0$, then $a=0$, or $b=0$

$2x-1=0$ or $3x^2-12x+5=0$

For $2x-1=0$, $\boxed{x=\frac{1}{2}=0.5}$

For $3x^2-12x+5=0$, we need to find Δ.

$\Delta = b^2-4ac = (-12)^2 - 4\cdot3\cdot5 = 84$, $x_{1,2} = \dfrac{-b \mp \sqrt{\Delta}}{2a}$, $x_{1,2} = \dfrac{6 \mp \sqrt{21}}{3}$

For finding y, we can solve $-0.096y^3 + 0.432y^2 - 0.352y + 0.08 = \dfrac{-2y+11}{15}$

Because y is common point for both intervals.

Multiplying both sides by 1000

$-96y^3 + 432y^2 - 352y + 80 = \dfrac{200(-2y+11)}{3}$

Multiplying both sides by 3

$-288y^3 + 1296y^2 - 1056y + 240 = 200(-2y+11)$

$-288y^3 + 1296y^2 - 1056y + 240 = -400y + 2200$

Organizing equation

$-288y^3 + 1296y^2 - 656y - 1960 = 0$

Organizing common factors

$-8(2y-5)(18y^2-36y-49) = 0$

Zero factor principle: If $a\cdot b=0$, then $a=0$ or $b=0$

$2y-5=0$ or $18y^2-36y-49=0$

For $2y-5=0$, $\boxed{y=\frac{5}{2}=2.5}$

For $18y^2-36y-49=0$, we need to find Δ.

$\Delta = b^2-4ac = (-36)^2 - 4\cdot18\cdot(-49) = 4824$, $x_{1,2} = \dfrac{-b \mp \sqrt{\Delta}}{2a}$, $x_{1,2} = \dfrac{6 \mp \sqrt{134}}{6}$

For t, we can see that from the graph, $f(y) = t$.

So, $f(y) = \frac{-2y+11}{15} = \frac{-2\cdot\frac{5}{2}+11}{15} = \frac{2}{5} = 0.4$, $\boxed{t = \frac{2}{5} = 0.4}$

We have chosen values $x = \frac{1}{2}$, $y = \frac{5}{2}$, $z = \frac{11}{2}$ and $t = \frac{2}{5}$.

For any pdf, area under its curve must be equal to 1.

So, we need to calculate the total area as $\int_x^y f(b)\,db + \frac{(z-y)\cdot t}{2} = 1$.

$$\int_x^y f(b)\,db = \int_{0.5}^{2.5} -0.096b^3 - 0.432b^2 - 0.352b + 0.08$$

$$= -\frac{0.096b^4}{4} - \frac{0.432b^3}{3} - \frac{0.352b^2}{2} + 0.08b \Bigg|_{0.5}^{2.5} = 0.4$$

For $\frac{(z-y)\cdot t}{2}$, $\frac{(5.5-2.5)\cdot 0.4}{2} = 0.6$

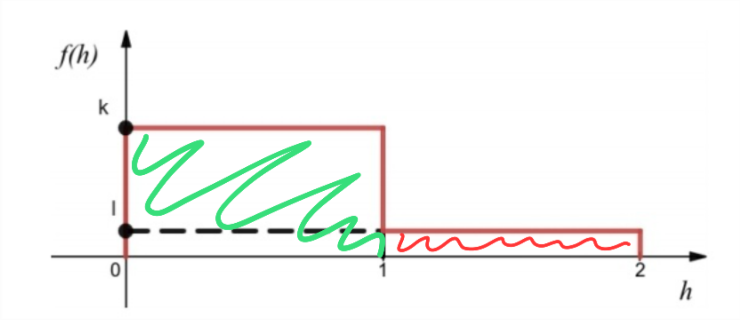
So, $\int_x^y f(b)\,db + \frac{(z-y)\cdot t}{2} = 0.4 + 0.6 = 1$, by this equality, we can say the values we have chosen are correct.

10. cdf of E given as $F(e;i,j) = (e-i)^2 - j$ for $(i+\sqrt{j}) \leq e \leq (i+\sqrt{j+1})$

In order to calculate pdf function of , we should derive cdf function because $F'(x) = f(x)$

$\frac{d}{de}(e-i)^2 - j = 2(e-i)$, So $\boxed{f(e) = 2(e-i)}$

13. For an pdf, area under its curve must be equal to 1.



We can calculate green shaded area as $(1-0) \cdot k$

Similarly, we can calculate red shaded area as $(2-1) \cdot l$

Since the total area is equal to 1, we can say $k + l = 1$ and $\boxed{l = 1 - k}$.

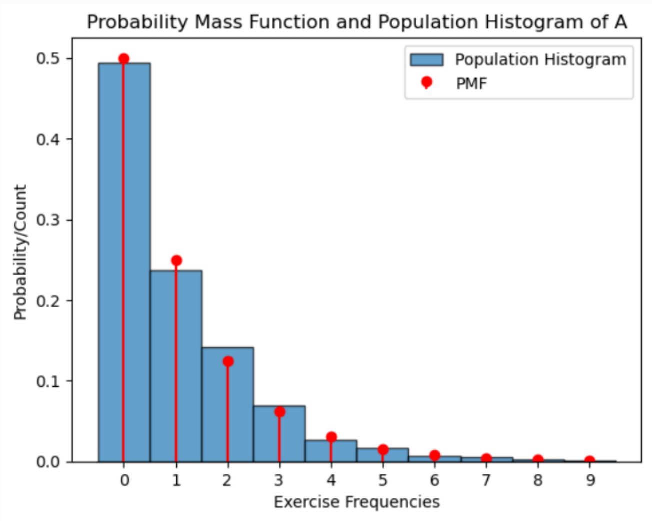
So, pdf becomes:

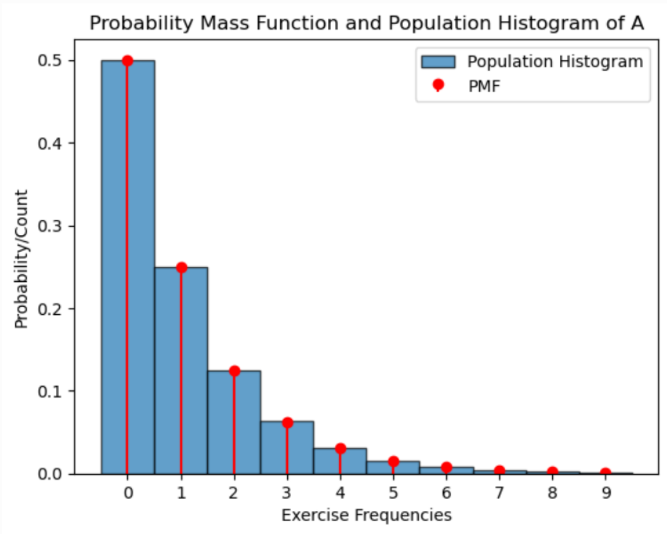
$f(h) = k \quad , 0 \leq h \leq 1$

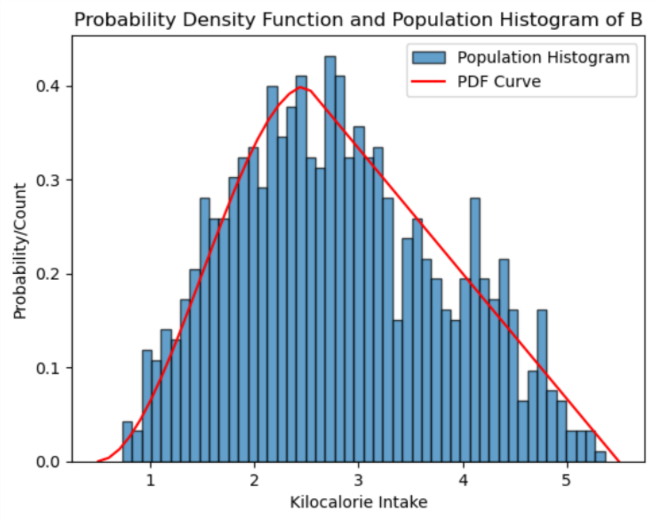
$f(h) = 1-k \quad , 1 < h \leq 2$

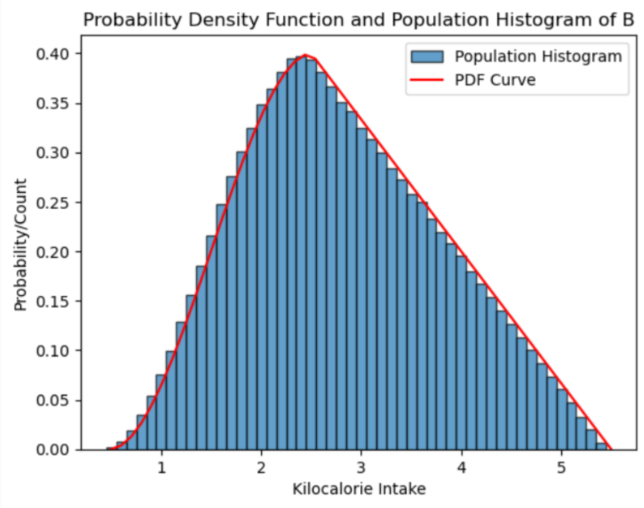
$f(h) = 0 \quad$, elsewhere

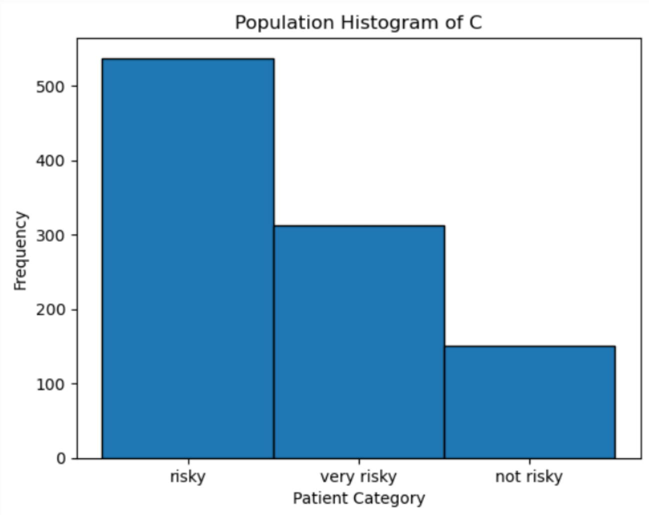
19.

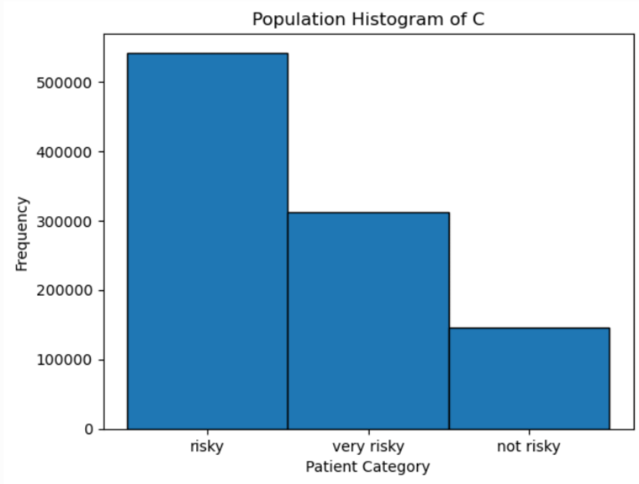


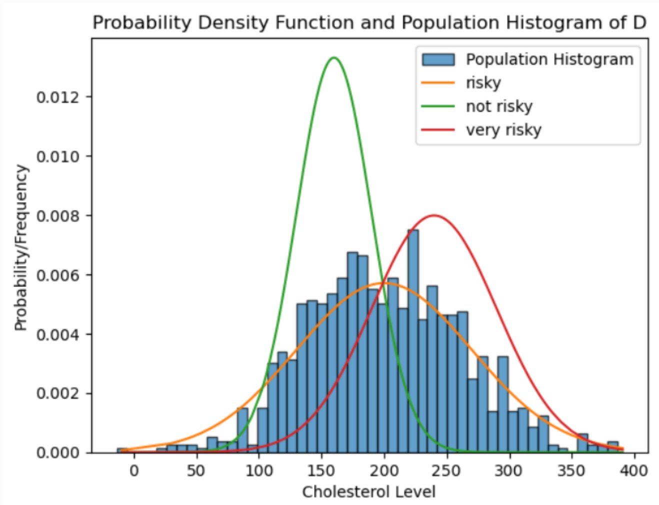


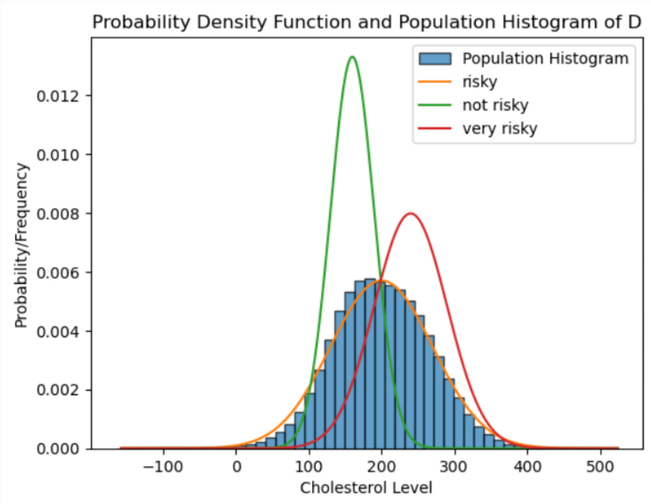

Probability Density Function and Population Histogram of B
Population Histogram
PDF Curve
Probability/Count
Kilocalorie Intake

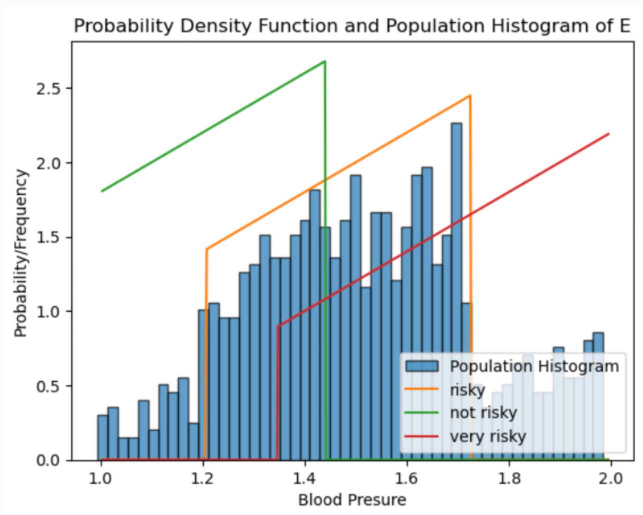

Probability Density Function and Population Histogram of B
Population Histogram
PDF Curve
Probability/Count
Kilocalorie Intake

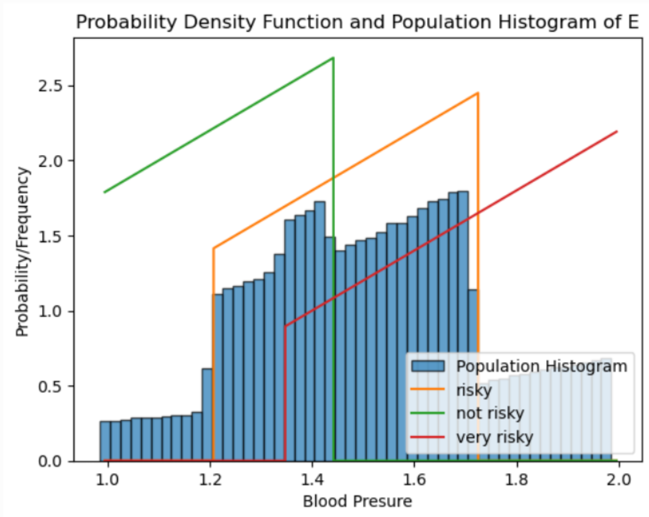

Population Histogram of C
Frequency
risky
very risky
not risky
Patient Category


Population Histogram of C
Frequency
risky
very risky
not risky
Patient Category


Probability Density Function and Population Histogram of D
Population Histogram
risky
not risky
very risky
Probability/Frequency
Cholesterol Level


Probability Density Function and Population Histogram of D
Population Histogram
risky
not risky
very risky
Probability/Frequency
Cholesterol Level


Probability Density Function and Population Histogram of E
Population Histogram
risky
not risky
very risky
Probability/Frequency
Blood Presure


Probability Density Function and Population Histogram of E
Population Histogram
risky
not risky
very risky
Probability/Frequency
Blood Presure

Task 2

6. We have found that in task 1, $A = X-1$ where X is a random variable of a geometric distribution. So, $E[A] = E[X-1]$.

Linear properties of expectations denotes that, for any random variables X and Y and any non-random numbers a, b, and c, we have:

$E[aX+bY+c] = aE[X] + bE[Y] + c$

In particular,

$E[X+Y] = E[X] + E[Y]$

$E[aX] = aE[X]$

$E[c] = c$

So, we can say that $E[X-1] = E[X] - 1$.

For geometric distribution, mean is equal to $1/P$, p is the probability of success.

In our case, $p = 0.5$. So, $E[X] = 2$ and $E[X-1] = \boxed{E[A] = 1}$

Similarly, we can calculate $Var[A] = Var[X-1]$.

Properties of variance denotes, for any random variable X and any non-random number b:

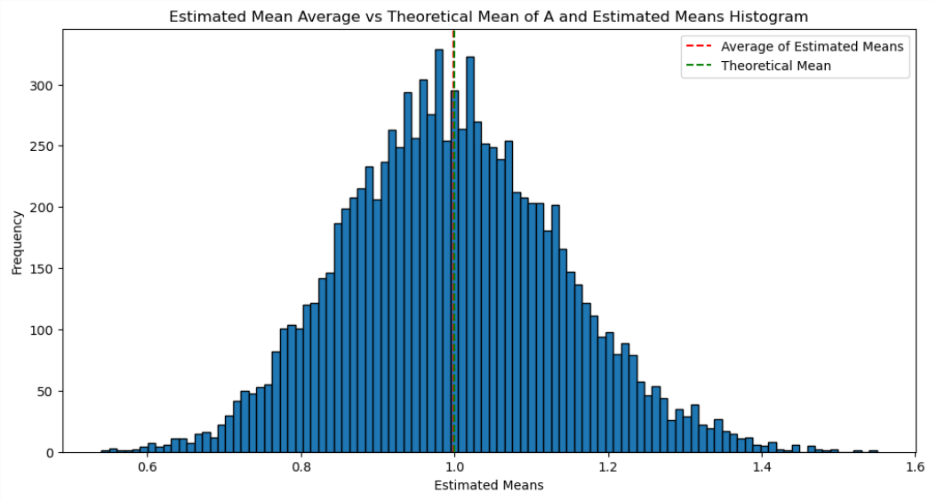
$Var[aX+b] = a^2 Var[X]$

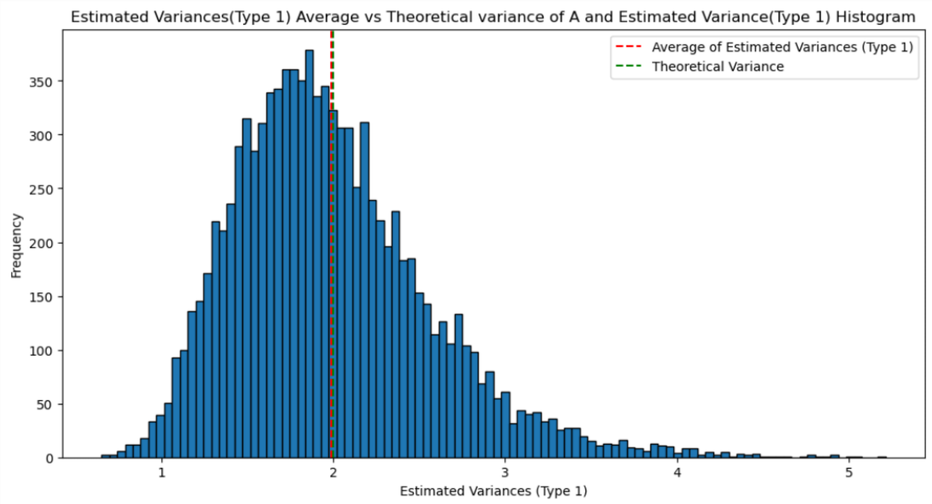
By this property, we can say $Var[X-1] = Var[X]$.

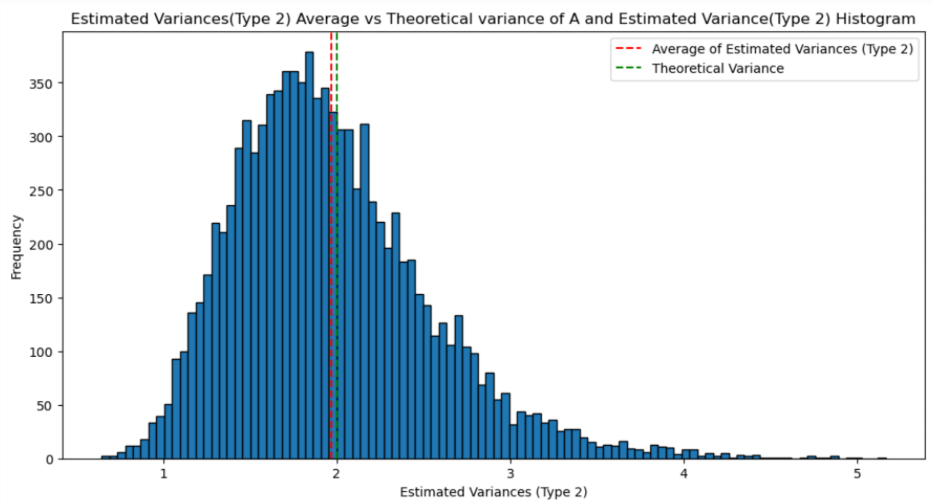
For geometric distribution, variance is equal to $(1-P)/p^2$, p is the probability of success.

In our case, $p = 0.5$. So, $Var[X] = \frac{1-0.5}{0.5^2} = 2$. $\boxed{Var[A] = 2}$

8.







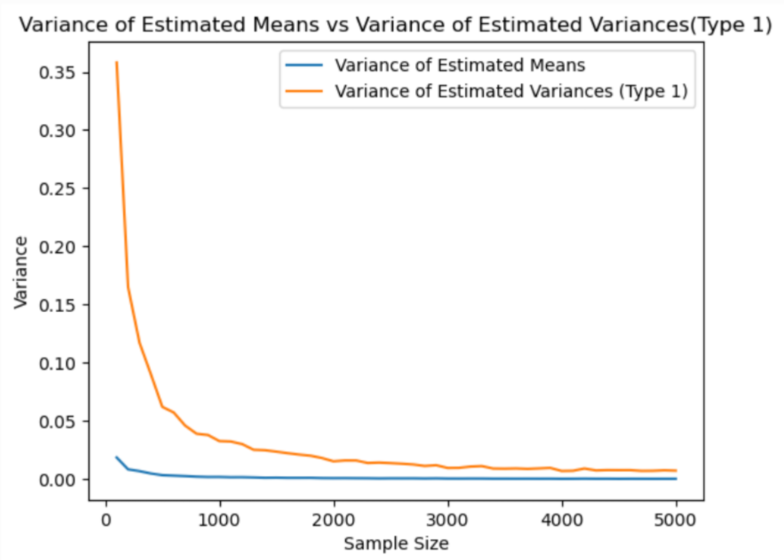
The graphs examine how simulation-derived estimated means and variances are distributed. A histogram illustrating the frequency of occurence of various estimated values appears in each figure, along with vertical lines displaying the average of the estimates and the theoretical mean / variance.

Histogram of estimated means offers perceptions on how accurately the estimating procedure captures the actual mean. In our figure, we see that, histogram is centered around the theoretical mean, and average of estimated means is closely aligned with theoretical mean. So, it indicates that estimation process performing well and estimates are unbiased.

For type 1 variances, where the denominator is (n-1), histogram is centered around the theoretical variance, and average of estimated variances is almost coincides with theoretical mean. So, it indicates that the estimations are unbiased and correctly represents the variability.

For type 2 variances, where the denominator is n, again histogram is centered around the theoretical variance, and average of estimated variances is slightly shifted to left of the theoretical variance with a very small difference. So, when it is compared to type 1, the estimation process of type 2 is slightly underestimating. The small difference, however, shows that the estimating method is still fairly accurate and reflects the overall variability.

10.



The estimated variances of the estimated means and the estimated variances of type 1 are examined in this figure with regard to sample size. We can see in graph, the variance of estimated variances is decreasing as sample size rises. This indicates larger samples lead to more reliable estimates with reduced variability. For variance of estimated means, the line do not show a clear decreasing, so, we can say that the sample size has already grown to the point where further increases have little effect on the accuracy of estimations.

Task 3

1. Let $\theta = \{\mu, \sigma\}$.

The problem can be framed as a conditional probability problem, where the goal is to maximize the probability that our data will be observed given θ. For a dataset of size n:
$f(x_1, x_2, \dots x_n | \theta)$.

We seek μ and σ values such that this probability density term is as high as it can be. By convention, X is typically the independent variable. However, since this is observed data and cannot change, we may take $x_1, x_2, \dots x_n$ as a constant in this situation and treat θ as the independent variable.

We can apply joint probability density as follows:

$$f(x_1, x_2, \dots, x_n | \theta) = f(x_1|\theta) \cdot f(x_2|\theta) \dots f(x_n|\theta) = \prod_i^n f(x_i|\theta)$$

Since we are looking for a maximum value, we should take a derivative with respect to θ and set to zero.

$$\frac{\partial}{\partial\theta} \prod_i^n f(x_i|\theta) = 0$$

We can take natural logarithm of both sides in order to make calculations easier. Natural logarithm is a monotonic function and it doesn't change the location of the maximum.

$$\frac{\partial}{\partial\theta} \prod_i^n f(x_i|\theta) \sim \frac{\partial}{\partial\theta} \ln \prod_i^n f(x_i|\theta) = \frac{\partial}{\partial\theta} \sum_i^n \ln(f(x_i|\theta)) = 0$$

Now, we apply this to normal distribution function:

$$\frac{\partial}{\partial\theta} \sum_i^n \ln\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\right) = 0$$

Using properties of natural logarithm, we can simplify this as:

$$\frac{\partial}{\partial\theta} \sum_i^n \ln\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\right) = \frac{\partial}{\partial\theta} \sum_i^n -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x_i-\mu)^2}{2\sigma^2}$$

Applying partial derivative with respect to μ:

$$\frac{\partial}{\partial \mu} \sum_{i}^{n} -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x_i-\mu)^2}{2\sigma^2} = -\frac{1}{2\sigma^2}\sum_{i}^{n} -2(x_i-\mu) = \frac{1}{\sigma^2}\sum_{i}^{n}(x_i-\mu)$$

Setting this term equal to zero, we get:

$$\frac{1}{\sigma^2}\sum_{i}^{n}(x_i-\mu) = 0 \qquad \sum_{i}^{n}(x_i-\mu) = 0 \qquad \boxed{\hat{\mu} = \frac{1}{n}\sum_{i}^{n} x_i}$$

Applying partial derivative with respect to σ:

$$\frac{\partial}{\partial \sigma} \sum_{i}^{n} -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x_i-\mu)^2}{2\sigma^2} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i}^{n}(x_i-\mu)^2$$

Setting this term equal to zero, we get:

$$-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i}^{n}(x_i-\mu)^2 = 0 \qquad \sum_{i}^{n}(x_i-\mu)^2 = n\sigma^2 \qquad \boxed{\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i}^{n}(x_i-\mu)^2}}$$

As conclusion, we found that these parameters work out to the exact same formulas we use for mean and standard deviation formulas.

3. In method of moments, we calculate the sample moment and population moment. Next, to choose such a member of this distribution whose properties are close to properties of our data, we match the moments. To estimate k parameters, we equate the first k population and sample moments.

k-th population moment: $\mu_k = E[X^k]$

k-th sample moment: $m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$

Since we have 2 parameters i and j, we should solve:

$\mu_1 = m_1$

$\mu_2 = m_2$

Calculating μ_1:

$$\mu_1 = \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e \cdot 2 \cdot (e-i)\, de = 2 \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e \cdot (e-i)\, de$$

organize the equation and apply sum rule

$$2 \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e \cdot (e-i)\, de = 2 \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 - ei\, de = 2\left(\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2\, de - \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} ei\, de \right)$$

$$\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2\, de = \frac{e^3}{3} \Bigg|_{i+\sqrt{j}}^{i+\sqrt{j+1}} = \frac{\sqrt{1+j}^{\,3} - \sqrt{j}^{\,3} - 3\sqrt{1+j} + 3\sqrt{j}}{3} + i$$

$$\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} ei\, de = \frac{ie^2}{2} \Bigg|_{i+\sqrt{j}}^{i+\sqrt{j+1}} = -\sqrt{1+j} + \sqrt{j} + \frac{i}{2}$$

$$2\left(\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2\, de - \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} ei\, de \right) = \frac{\sqrt{1+j}^{\,3} - \sqrt{j}^{\,3} - 3\sqrt{1+j} + 3\sqrt{j}}{3} + i - \left(-\sqrt{1+j} + \sqrt{j} + \frac{i}{2} \right)$$

Simplifying equation we get:

$$\mu_1 = \frac{2(i+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3} + i$$

From $\mu_1 = m_1$, we can get:

$$\boxed{i = m_1 - \frac{2(1+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3}}$$

Calculating μ_2:

$$\mu_2 = \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 \cdot 2 \cdot (e-i)\, de = 2\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 \cdot (e-i)\, de$$

organize the equation and apply sum rule

$$2\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 \cdot (e-i)\, de = 2\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^3 - e^2 i\, de = 2\left(\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^3\, de - \int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 i\, de\right)$$

$$\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^3\, de = \frac{e^4}{4}\Bigg|_{i+\sqrt{j}}^{i+\sqrt{j+1}} = \frac{-4i^3\sqrt{j} + 4i^3\sqrt{j+1} + 6i^2 - 4i\sqrt{j}^3 + 4i\sqrt{j+1}^3 + 2j + 1}{4}$$

$$\int_{i+\sqrt{j}}^{i+\sqrt{j+1}} e^2 i\, de = \frac{ie^3}{3}\Bigg|_{i+\sqrt{j}}^{i+\sqrt{j+1}} = \frac{i\left(-3i^2\sqrt{j} + 3i^2\sqrt{j+1} + 3i - \sqrt{j}^3 + \sqrt{j+1}^3\right)}{3}$$

Simplifying equation we get:

$$\mu_2 = i^2 + \frac{4}{3} i \left(-j^{\frac{3}{2}} + (j+1)^{\frac{3}{2}}\right) + j + \frac{1}{2}$$

By substituting i, we get:

$$\boxed{\mu_2 = \left(m_1 - \frac{2(1+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3}\right)^2 + j + \frac{1}{2} + \frac{4}{3}\left(m_1 - \frac{2(1+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3}\right)\left(-j^{\frac{3}{2}} + (j+1)^{\frac{3}{2}}\right)}$$

So, $f(j) = \mu_2 - m_2 = 0$

$$\boxed{f(j) = \left(m_1 - \frac{2(1+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3}\right)^2 + j + \frac{1}{2} + \frac{4}{3}\left(m_1 - \frac{2(1+j)^{\frac{3}{2}} - 2j^{\frac{3}{2}}}{3}\right)\left(-j^{\frac{3}{2}} + (j+1)^{\frac{3}{2}}\right) - m_2}$$

5. $f(h) = \begin{cases} k, & 0 \leq h \leq 1 \\ 1-k, & 1 < h \leq 2 \end{cases}$

We can write the likelihood function as:

$$L(k) = \left(\prod_{0 \leq h \leq 1} k\right)\left(\prod_{1 < h \leq 2} (1-k)\right), \quad k \in (0,1)$$

This is equal to:

$L(k) = k^n (1-k)^m$, where n is number of samples in range $0 \leq h \leq 1$
and m is number of samples in range $1 < h \leq 2$

Applying natural logarithm to both sides:

$$\ln(L(k)) = n \cdot \ln(k) + m \ln(1-k)$$

Taking the derivative and setting to 0:

$$\frac{\partial}{\partial k} \ln(L(k)) = \frac{n}{k} - \frac{m}{1-k} = 0$$

Simplifying this equation we get:

$$n(1-k) - mk = 0$$

$$n - nk - mk = 0$$

$$(n+m)k = n$$

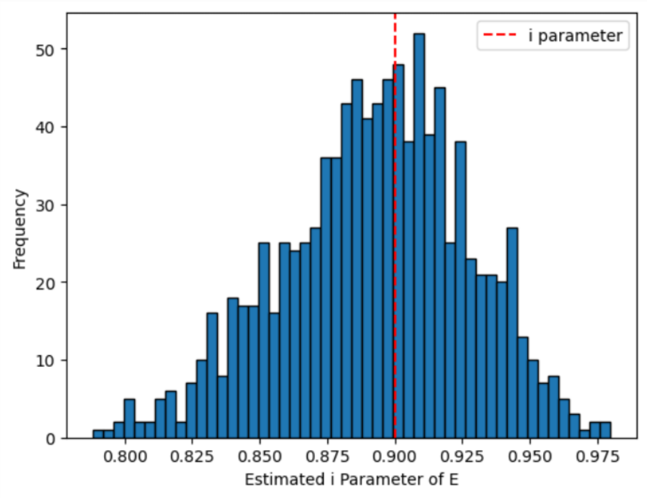
$$\boxed{\hat{k} = \frac{n}{n+m}}$$

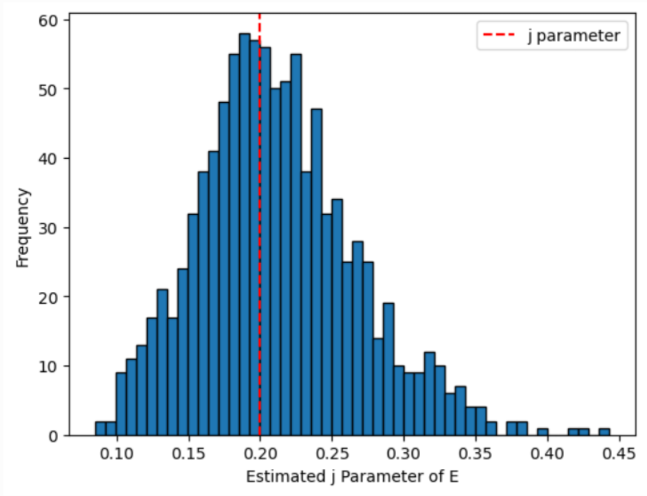
Now, substituting $l = 1-k$ we get

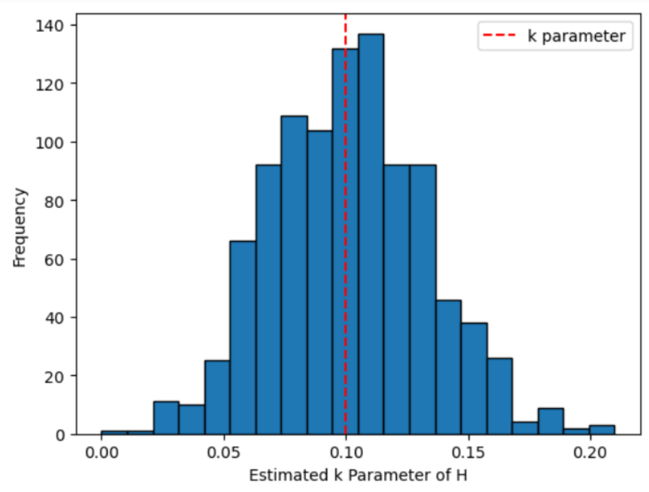
$$\hat{l} = 1 - \frac{n}{n+m}$$

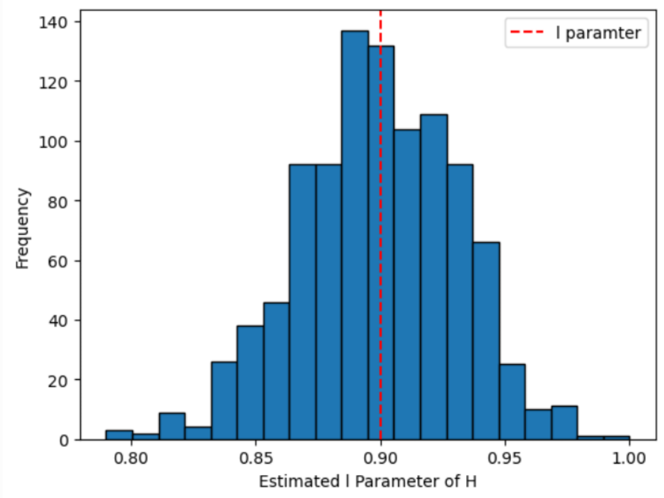
$$\boxed{\hat{l} = \frac{m}{n+m}}$$

9.


i parameter
Frequency
Estimated i Parameter of E

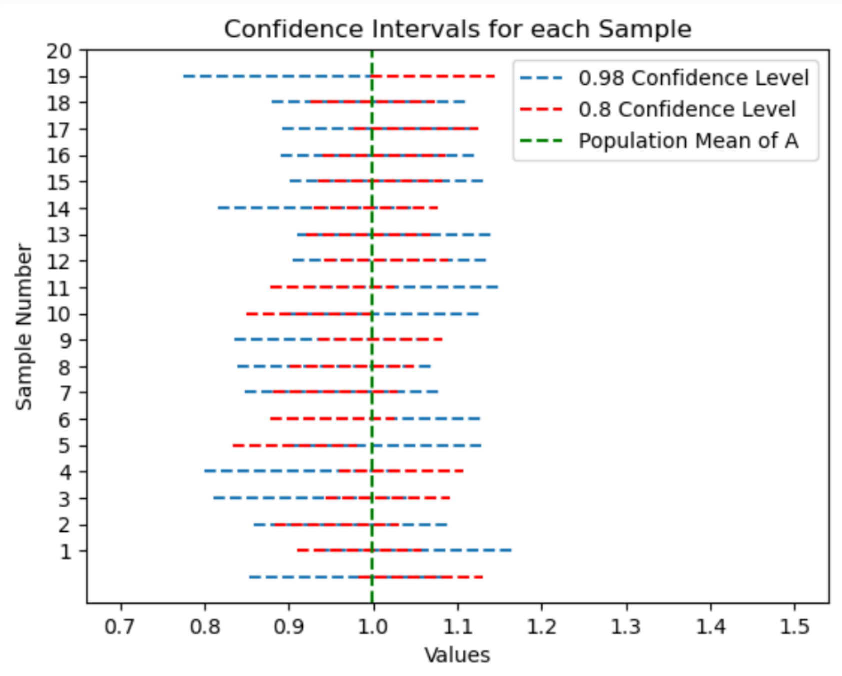

j parameter
Frequency
Estimated j Parameter of E


k parameter
Frequency
Estimated k Parameter of H


l paramter
Frequency
Estimated l Parameter of H

Task 4

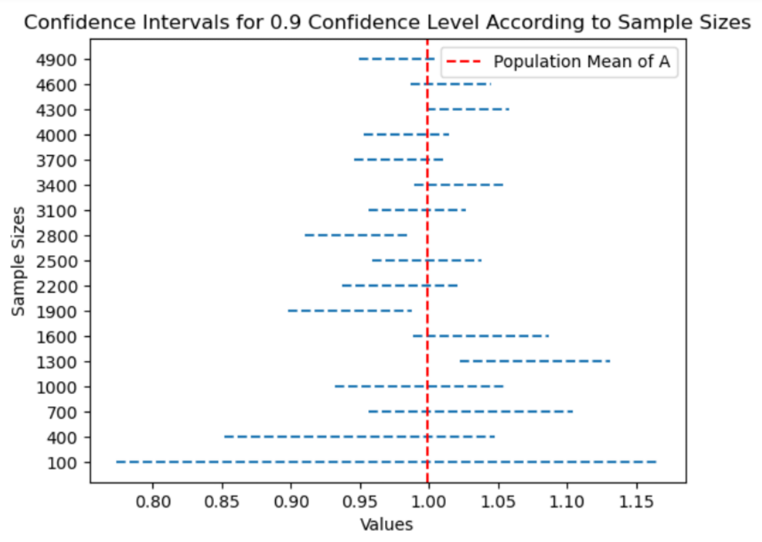
3.


Confidence Intervals for each Sample
0.98 Confidence Level
0.8 Confidence Level
Population Mean of A
Sample Number
Values

A confidence interval, which is build around a point estimate, identifies the range in which the real population parameter is most likely to fall. The width of the confidince interval depends on several factors, including the desired level of confidence.

According to the graphs, we can observe from a comparison of two confidence intervals that 98% confidence interval is larger than 80% confidence interval. This is expected because a wider interval is needed to achieve a higher level of certainty at a higher confidence level. Compared to the 80% confidence interval, the 98% confidence interval offers a higher level of certainity. It is more likely that the true population parameter falls inside the %98 confidence interval's larger range. Contrarily, the 80% confidence interval is more restricted offering less certainity but a more accurate prediction.

5.



Confidence intervals tend to get smaller as sample size grows. This is due to the fact that larger samples offer more information and reduce the uncertainity involved in predicting population parameters and provides more precise estimates. With a larger sample size, margin of error $\pm z \cdot \frac{\sigma}{\sqrt{n}}$ becomes smaller, resulting in a narrower confidence interval.

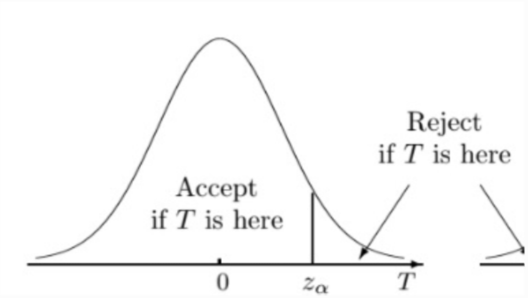
Task 5

Let μ be the mean after the campaign and μ_0 be the mean before the campaign.

$H_0 : \mu = \mu_0$

$H_A : \mu > \mu_0$

Since we are testing whether the mean has increased, this is a one-tail test. (Right-tail)



T: test statistics

For right-tail, we accept null hypothesis if test statistics smaller than critical value. Otherwise, we reject null hypothesis in favor of H_A.

Step 1: Test statistics. We are given $\sigma = \sqrt{2}$, $n = 500$, $\alpha = 0.03$, $\mu_0 = 1$ and $\bar{x} = 1.2$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{1.2 - 1}{\sqrt{2} / \sqrt{500}} \approx 3.16$$

Step 2: Acceptance and rejection regions. The critical value is

$$z_\alpha = z_{0.03} \approx 1.8808$$

With the right-tail alternative, we

reject H_0 if $z \geqslant 1.8808$

accept H_0 if $z < 1.8808$

Step 3: Result. Our test statistics $z = 3.16$ belongs to rejection region; therefore, we reject the null hypothesis. This means that, at a 3% significance level, there is sufficient evidence to conclude that the mean of number of times a patient engages in physical exercise each week has increased after the campaign.

Task 6

2. Bayes Theorem determines the probability of an event occuring given the probability of another event occuring.

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)}$$

Now, if any two events X and Y are independent, then

$$P(X,Y) = P(X) \cdot P(Y)$$

Hence we reach to the result:

$$P(X|y_1,\dots,y_n) = \frac{P(y_1|X)\, P(y_2|X) \dots P(y_n|X) \cdot P(X)}{P(y_1) \cdot P(y_2) \dots P(y_n)}$$

For our problem, we need to find posterior probability of each patient category (C) given D, E, H.

$$\boxed{P(C|D,E,H) = \frac{f_{D|C}(D|C) \cdot f_{E|C}(E|C) \cdot f_{H|C}(H|C)}{f(D) \cdot f(E) \cdot f(H)} \cdot P(C)}$$