Problem 1

(a) *Express $P(X_i = 1)$ as a function of a and π.*

We have

$$\begin{aligned} P(X_i = 1) &= P(X_i = 1, \text{Group} = A) + P(X_i = 1, \text{Group} = B) \\ &= P(X_i = 1 \mid \text{Group} = A) \cdot P(\text{Group} = A) + P(X_i = 1 \mid \text{Group} = B) \cdot P(\text{Group} = B) \\ &= a \cdot P(\text{Group} = A) + (1 - a) \cdot P(\text{Group} = B) \\ &= a \cdot \pi + (1 - a) \cdot (1 - \pi) \ . \end{aligned}$$

Naturally, we have that

$$\begin{aligned} P(X_i = 0) &= P(X_i = 0, \text{Group} = A) + P(X_i = 0, \text{Group} = B) \\ &= P(X_i = 0 \mid \text{Group} = A) \cdot P(\text{Group} = A) + P(X_i = 0 \mid \text{Group} = B) \cdot P(\text{Group} = B) \\ &= (1 - a) \cdot P(\text{Group} = A) + a \cdot P(\text{Group} = B) \\ &= (1 - a) \cdot \pi + a \cdot (1 - \pi) \ . \end{aligned}$$

(Optional) Because the random variable X_i can only have 2 potential values, we can now easily check whether our expressions are correct as $P(X_i = 1) + P(X_i = 0) = 1$. We have

$$\begin{aligned} P(X_i = 1) + P(X_i = 0) &= a \cdot \pi + (1 - a) \cdot (1 - \pi) + (1 - a) \cdot \pi + a \cdot (1 - \pi) \\ &= a \cdot \pi + (1 - \pi) - a + \pi \cdot a + \pi - \pi \cdot a + a - \pi \cdot a \\ &= a \cdot \pi + 1 - \pi + \pi - \pi \cdot a \\ &= 1 \ . \end{aligned}$$

(b) *Give expressions for the joint probability of the observed data $X_1, X_2, \ldots, X_n$, and the corresponding likelihood function and loglikelihood function.*

We aim to express the computation for

$$P(X_1, X_2, \ldots, X_n) \ .$$

Because we know that the n samples are drawn with the i.i.d. assumption, it holds that

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i) \ .$$

Hence,

$$
\begin{aligned}
P(X_1, X_2, \dots, X_n) &= \prod_{i=1}^{n} P(X_i) \\
&= \prod_{i=1}^{n} \begin{cases} P(X_i = 1) & \text{if } X_i = 1 \\ P(X_i = 0) & \text{if } X_i = 0 \end{cases} \\
&= P(X_i = 1)^m \cdot P(X_i = 0)^{n-m} \\
&= (a \cdot \pi + (1-a) \cdot (1-\pi))^m \cdot ((1-a) \cdot \pi + a \cdot (1-\pi))^{n-m} \qquad (1)
\end{aligned}
$$

which is the likelihood function of the observed samples.

Let us now derive the experession for the log likelihood of the data:

$$
\begin{aligned}
\log P(X_1, X_2, \dots, X_n) &= \log \left[(a \cdot \pi + (1-a) \cdot (1-\pi))^m \cdot ((1-a) \cdot \pi + a \cdot (1-\pi))^{n-m} \right] \\
&= m \cdot \log \left[a \cdot \pi + (1-a) \cdot (1-\pi) \right] + (n-m) \cdot \log \left[(1-a) \cdot \pi + a \cdot (1-\pi) \right] \qquad (2)
\end{aligned}
$$

(c) *Derive an expression for the maximum likelihood estimator* π_{ML} *of* π.

We now aim to estimate π. To this end, we replace π with π_{ML} in Equation 2 and derive the log likelihood w.r.t. π_{ML}. We have

$$
\begin{aligned}
&\frac{\partial}{\partial \pi_{ML}} \log P(X_1, X_2, \dots, X_n) \\
=&\frac{\partial}{\partial \pi_{ML}} m \cdot \log \left[a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML}) \right] + (n-m) \cdot \log \left[(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML}) \right] \\
=&m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} \cdot (a - (1-a)) + (n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \cdot ((1-a) - a) \,.
\end{aligned}
$$

We can now find the critical points by setting the gradient to zero:

$$
\begin{aligned}
\frac{\partial}{\partial \pi_{ML}} \log P(X_1, X_2, \dots, X_n) &\stackrel{!}{=} 0 \\
m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} \cdot (a - (1-a)) + (n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \cdot ((1-a) - a) &= 0 \qquad (3)
\end{aligned}
$$

where we observe that the the equation holds whenever $(1-a) = a$, because then both additive terms are zero. It is easy to see that $a = 0.5$ is the only real valued solution for this condition. However, as we are interested in the maximum likelihood estimator of π_{ML}, setting $a = 0.5$ prohibits us from finding the maximum value.

Explain why setting $a = 0.5$ is problematic in more details...

Therefore, we have to consider the case where $a \neq 0.5$.

$$
\begin{aligned}
m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} \cdot (a - (1-a)) &= -(n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \cdot ((1-a) - a) \\
m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} \cdot (a - (1-a)) &= (n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \cdot (a - (1-a)) \\
m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} &= (n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \\
m \cdot \frac{1}{(2a-1) \cdot \pi_{ML} + 1 - a} &= (n-m) \cdot \frac{1}{(1-2a) \cdot \pi_{ML} + a} \\
m \cdot \frac{1}{(2a-1) \cdot \pi_{ML} + 1 - a} &= (n-m) \cdot \frac{1}{a - (2a-1) \cdot \pi_{ML}} \\
m \cdot (a - (2a-1) \cdot \pi_{ML}) &= (n-m) \cdot ((2a-1) \cdot \pi_{ML} + 1 - a) \\
am - (2a-1) \cdot m \cdot \pi_{ML} &= (n-m) \cdot (2a-1) \cdot \pi_{ML} + (n-m) \cdot (1-a) \\
am - (n-m) \cdot (1-a) &= (n-m) \cdot (2a-1) \cdot \pi_{ML} + (2a-1) \cdot m \cdot \pi_{ML} \\
am - (n-m) \cdot (1-a) &= \pi_{ML} \cdot (2a-1)((n-m) + m) \\
am - (n-m) + a \cdot n - am &= \pi_{ML} \cdot (2a-1)(n) \\
n \cdot (a-1) + m &= \pi_{ML} \cdot (2a-1)(n) \\
\frac{n \cdot (a-1) + m}{(2a-1)(n)} &= \pi_{ML} \\
\frac{n \cdot (a-1) + m}{n} \cdot \frac{1}{2a-1} &= \pi_{ML} \\
\left(a - 1 + \frac{m}{n}\right) \cdot \frac{1}{2a-1} &= \pi_{ML} \qquad (4)
\end{aligned}
$$

We verify that this is indeed a maximum by plugging the result from Equation 4 into the second derivative:

$$
\begin{aligned}
&\frac{\partial^2}{\partial {\pi_{ML}}^2} \log P(X_1, X_2, \dots, X_n) \\
=&\frac{\partial}{\partial \pi_{ML}} m \cdot \frac{1}{a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML})} \cdot (a - (1-a)) + (n-m) \cdot \frac{1}{(1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML})} \cdot ((1-a) - a) \\
=&m \cdot (a - (1-a)) \cdot \frac{1}{(a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML}))^2} \cdot (a - (1-a)) \\
&+ (n-m) \cdot ((1-a) - a) \cdot \frac{1}{((1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML}))^2} \cdot ((1-a) - a) \\
=&2m \cdot (a - (1-a)) \cdot \frac{1}{(a \cdot \pi_{ML} + (1-a) \cdot (1-\pi_{ML}))^2} + 2(n-m) \cdot ((1-a) - a) \cdot \frac{1}{((1-a) \cdot \pi_{ML} + a \cdot (1-\pi_{ML}))^2}
\end{aligned}
$$

This will be loooong I think...

(d) *Can you think of an easier way to arrive at the same estimator as the maximum likelihood estimator?*

We observe that X_i follows a Bernoulli distribution. We now first show that this holds.

Recall that the definition of a Bernoulli distribution is given by

$$
\begin{aligned}
P(X=1) &= p \\
P(X=0) &= 1-p \; .
\end{aligned}
$$

Therefore, we now aim to show that $P(X_i)$ follows a Bernoulli distribution. From (a), we have that We have

$$P(X_i = 1) = a \cdot \pi + (1-a) \cdot (1-\pi)$$

and

$$P(X_i = 0) = (1-a) \cdot \pi + a \cdot (1-\pi) \; .$$

We multiply out the expression to get

$$
\begin{aligned}
P(X_i = 1) &= a \cdot \pi + 1 - \pi - a \cdot (1-\pi) \\
&= a \cdot \pi + 1 - \pi - a + a \cdot \pi \\
&= 2a \cdot \pi + 1 - \pi - a \\
&= 1 + 2a \cdot \pi - \pi - a \qquad (5) \\
P(X_i = 0) &= (1-a) \cdot \pi + a \cdot (1-\pi) \\
&= \pi - a \cdot \pi + a - a \cdot \pi \\
&= -2a \cdot \pi + \pi + a \; . \qquad (6)
\end{aligned}
$$

Therefore, with $p = 1 + 2a \cdot \pi - \pi - a$ (from Equation 5), we have

$$
\begin{aligned}
P(X=0) &= 1 - p \\
&= 1 - P(X=1) \\
&= 1 - (1 + 2a \cdot \pi - \pi - a) \\
&= 1 - 1 - 2a \cdot \pi + \pi + a \\
&= -2a \cdot \pi + \pi + a \qquad (7)
\end{aligned}
$$

where Equation 7 matches the expression in Equation 6. Therefore, we have shown that X_i follows a Bernoulli distribution with parameter $p = 1 + 2a \cdot \pi - \pi - a$.

Following from this result, because the random variable m is the sum of n i.i.d. Bernoulli random variables, we know that m follows a Binomial distribution with parameters n and p. Hence, the likelihood of observing m ones in n samples is given by the probability density function of the Binomial distribution:

$$\Pr(X = m) = \binom{n}{m} p^m (1-p)^{n-m} \; . \qquad (8)$$

While Equation 8 differs from the likelihood function we derived in Equation 1, we can see that the $\binom{n}{m}$ term cancels out when deriving the log likelihood; First, the log likelihood transforms the multiplication into a sum with a log term, and then when deriving w.r.t. π_{ML}, the log $\binom{n}{m}$ becomes zero and therefore disappears. Hence, we end up with the same expression for the log likelihood.

As the Binomial distribution is well-known, we know that the maximum likelihood estimator of p of the Binomial distribution is given by $p_{ML} = \frac{m}{n}$, which is the proportion of successes, i.e., the number of $X_i = 1$ in n samples.[1]

Finally, to derive the maximum likelihood estimator of π_{ML}, we equate p_{ML} and $P(X_i = 1)$ (as $P(X_i = 1) = p$), and then solve for π_{ML}:

$$\begin{aligned} P(X_i = 1) &= p_{ML} \\ 1 + 2a \cdot \pi_{ML} - \pi_{ML} - a &= \frac{m}{n} \\ 2a \cdot \pi_{ML} - \pi_{ML} &= \frac{m}{n} + a - 1 \\ \pi_{ML} \cdot (2a - 1) &= \frac{m}{n} + a - 1 \\ \pi_{ML} &= \frac{\frac{m}{n} + a - 1}{(2a-1)} \; . \end{aligned} \tag{9}$$

We clearly observe that Equation 9 matches Equation 4, hence, we found the same maximum likelihood estimator of π_{ML} as in (c).

(e) *Show that π_{ML} is an unbiased estimator of π.*

We aim to show that $\mathbb{E}[\pi_{ML}] = \pi$. To this end, we will make use of the knowledge introduced in (d), namely that X_i follows a Bernoulli distribution.

We have

$$\begin{aligned} \mathbb{E}\left[\pi_{ML}\right] &= \mathbb{E}\left[\frac{\frac{m}{n} + a - 1}{(2a-1)}\right] \\ &= \mathbb{E}\left[\frac{\frac{\sum_{i=1}^{n} X_i}{n} + a - 1}{(2a-1)}\right] \\ &= \frac{\sum_{i=1}^{n} \mathbb{E}[X_i] + a \cdot n - 1 \cdot n}{(2a-1) \cdot n} \; . \end{aligned} \tag{10}$$

We now compute the expectation of X_i:

$$\begin{aligned} \mathbb{E}[X_i] &= P(X_i = 1) \cdot 1 + P(X_i = 0) \cdot 0 \\ &= P(X_i = 1) \\ &= a \cdot \pi + (1 - a) \cdot (1 - \pi) \\ &= (1 - a) + (2a - 1) \cdot \pi \; . \end{aligned} \tag{11}$$

[1] See `https://en.wikipedia.org/wiki/Binomial_distribution#Estimation_of_parameters`.

Finally, we can substitute Equation 11 into Equation 10 to get

$$\begin{aligned}
\mathbb{E}\left[\pi_{ML}\right] &= \frac{\sum_{i=1}^{n} \mathbb{E}[X_i] + a \cdot n - 1 \cdot n}{(2a-1) \cdot n} \\
&= \frac{\sum_{i=1}^{n} \left((1-a) + (2a-1) \cdot \pi\right) + a \cdot n - 1 \cdot n}{(2a-1) \cdot n} \\
&= \frac{n \cdot \left((1-a) + (2a-1) \cdot \pi\right) + a \cdot n - 1 \cdot n}{(2a-1) \cdot n} \\
&= \frac{\left((1-a) + (2a-1) \cdot \pi\right) + a - 1}{(2a-1)} \\
&= \frac{-(a-1) + (2a-1) \cdot \pi + (a-1)}{(2a-1)} \\
&= \frac{(2a-1) \cdot \pi}{(2a-1)} \\
&= \pi \ .
\end{aligned} \tag{12}$$

We have shown that the expected value of π_{ML} is equal to the true value of π (Equation 12). Therefore, we can conclude that our estimator for π_{ML} is unbiased.

(f) *Derive an expression for the variance of π_{ML}. Analyse its dependence on a.*

We have

$$\begin{aligned}
\mathbb{V}[\pi_{ML}] &= \mathbb{V}\left[\frac{\frac{m}{n} + a - 1}{(2a-1)}\right] \\
&= \mathbb{V}\left[\frac{m}{n} + a - 1\right] \cdot \frac{1}{(2a-1)^2} \\
&= \mathbb{V}\left[\frac{1}{n}\left(m + a \cdot n - 1 \cdot n\right)\right] \cdot \frac{1}{(2a-1)^2} \\
&= \mathbb{V}\left[m + a \cdot n - 1 \cdot n\right] \cdot \frac{1}{(2a-1)^2 \cdot n^2} \\
&= \mathbb{V}\left[\sum_{i=1}^{n} X_i + a \cdot n - 1 \cdot n\right] \cdot \frac{1}{(2a-1)^2 \cdot n^2} \\
&= \mathbb{V}\left[\sum_{i=1}^{n} X_i\right] \cdot \frac{1}{(2a-1)^2 \cdot n^2} \\
&= \sum_{i=1}^{n} \mathbb{V}\left[X_i\right] \cdot \frac{1}{(2a-1)^2 \cdot n^2} \ .
\end{aligned} \tag{13}$$

We now compute the variance of X_i. We have:

$$\begin{aligned}
\mathbb{E}[X_i^2] &= P(X_i = 1) \cdot 1^2 + P(X_i = 0) \cdot 0^2 \\
&= P(X_i = 1) \\
&= (1-a) + (2a-1) \cdot \pi \ ,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[X_i]^2 &= ((1-a) + (2a-1)\cdot\pi)^2 \\
&= (1-a)^2 + (2a-1)^2\cdot\pi^2 + 2\cdot(1-a)\cdot(2a-1)\cdot\pi \\
&= 1^2 + a^2 - 2a + 4a^2\pi^2 + \pi^2 - 4a\cdot\pi^2 + (2-2a)\cdot(2a-1)\cdot\pi \\
&= 1 + a^2 - 2a + 4a^2\pi^2 + \pi^2 - 4a\cdot\pi^2 + (4a - 2 - 4a^2 + 2a)\cdot\pi \\
&= 1 + \pi\left(6a - 4a^2 - 2\right) + \pi^2(4a^2 - 4a + 1) - 2a + a^2\ .
\end{aligned}$$

Therefore, the variance of X_i is given by

$$\begin{aligned}
V(X_i) &= \mathbb{E}\left[X_i^2\right] - E[X_i]^2 \\
&= \pi(2a-1) + 1 - a - \left(1 + \pi\left(6a - 4a^2 - 2\right) + \pi^2(4a^2 - 4a + 1) - 2a + a^2\right) \\
&= \pi(2a-1) + 1 - a - 1 - \pi\left(6a - 4a^2 - 2\right) - \pi^2(4a^2 - 4a + 1) + 2a - a^2 \\
&= \pi(2a-1) - \pi\left(6a - 4a^2 - 2\right) - \pi^2(4a^2 - 4a + 1) + a - a^2 \\
&= \pi\left(2a - 1 - (6a - 4a^2 - 2)\right) - \pi^2(4a^2 - 4a + 1) + a - a^2 \\
&= \pi\left(2a - 1 - 6a + 4a^2 + 2\right) - \pi^2(4a^2 - 4a + 1) + a - a^2 \\
&= \pi\left(1 - 4a + 4a^2\right) - \pi^2(1 - 4a + 4a^2) + a - a^2 \\
&= \left(1 - 4a + 4a^2\right)\cdot(\pi - \pi^2) + a - a^2\ . \qquad (14)
\end{aligned}$$

Finally, to compute the variance of the estimator of π_{ML}, we plug Equation 14 into Equation 13 and get

$$\begin{aligned}
\mathbb{V}[\pi_{ML}] &= \sum_{i=1}^{n} \mathbb{V}\left[X_i\right]\cdot\frac{1}{(2a-1)^2\cdot n^2} \\
&= \sum_{i=1}^{n}\left(\left(1 - 4a + 4a^2\right)\cdot(\pi - \pi^2) + a - a^2\right)\cdot\frac{1}{(2a-1)^2\cdot n^2} \\
&= n\cdot\left(\left(1 - 4a + 4a^2\right)\cdot(\pi - \pi^2) + a - a^2\right)\cdot\frac{1}{(2a-1)^2\cdot n^2} \\
&= \left(\left(1 - 4a + 4a^2\right)\cdot(\pi - \pi^2) + a - a^2\right)\cdot\frac{1}{(2a-1)^2\cdot n} \\
&= \left((2a-1)^2\cdot(\pi - \pi^2) + a - a^2\right)\cdot\frac{1}{(2a-1)^2\cdot n} \\
&= \frac{(2a-1)^2\cdot(\pi - \pi^2)}{(2a-1)^2\cdot n} + \frac{a - a^2}{(2a-1)^2\cdot n} \\
&= \frac{\pi - \pi^2}{n} + \frac{a - a^2}{(2a-1)^2\cdot n}\ .
\end{aligned}$$

We correctly observe that the variance is undefined for $a = 0.5$. Furthermore, we observe that the variance decreases as n increases, which is to be expected as we get more evidence with more data. Lastly, we observe that whenever a approaches either 0 or 1, the variance decreases. The same phenomenon is observed for the true value of π. Figure 1 shows the log-scaled variance of π_{ML} as a function of a and π, which clearly displays the behavior previously described.

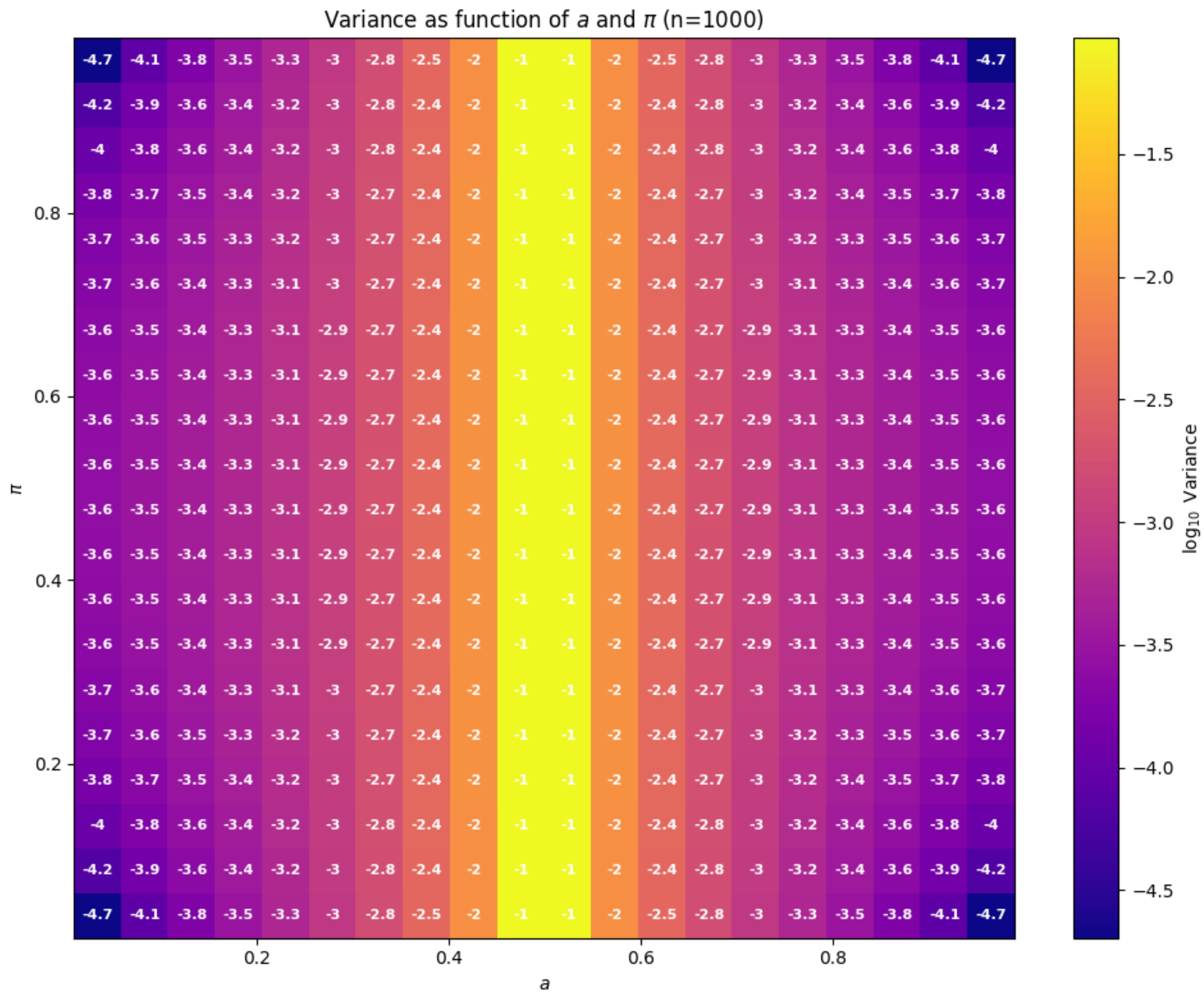

Figure 1: Variance of π_{ML} as function of a and π for different values of n. Here, the values are log-scaled for better display. Note that we restrict $a \in [0.01, 0.99]$ and $\pi \in [0.01, 0.99]$ for the plot. Hence, very low values are expected whenever a or π exceeds these bounds.

(g) *Give an expression for the bias of this naive estimator. Is it asymptotically unbiased?*

We aim to recover the value of π from direct questioning. However, persons belonging to the sensitive group will lie with probability ℓ, whereas persons belonging to the non-sensitive group will answer truthfully and therefore lie with probability 0.

Naivily, we could estimate the value of π as the fraction of persons answering yes to the question:

$$\pi_{naive} = \frac{\sum_{i=1}^{n} X_i}{n} \, .$$

We can model the whether a person belongs to the sensitive group as a hidden variable $S_i \in \{0, 1\}$ with $S_i = 1$ describing that the person belongs to the sensitive group, and $S_i = 0$ describes that a person belongs to the non-sensitive group.

Naturally, we have that

$$\begin{aligned} P(S_i = 1) &= \pi \\ P(S_i = 0) &= 1 - \pi \end{aligned}$$

which is distributed according to a Bernoulli distribution with parameter $p = \pi$.

We can now express the probability of a person answering yes or no to the question by marginalizing over the hidden variable S_i:

$$\begin{aligned} P(X_i = 1) &= P(X_i = 1, S_i = 1) + P(X_i = 1, S_i = 0) \\ &= \pi \cdot (1 - \ell) + 0 \\ P(X_i = 0) &= P(X_i = 0, S_i = 1) + P(X_i = 0, S_i = 0) \\ &= \pi \cdot \ell + (1 - \pi) \ . \end{aligned}$$

We verify that the probabilities are correct by adding both $P(X_i = 1)$ and $P(X_i = 0)$ and getting 1:

$$\begin{aligned} P(X_i = 1) + P(X_i = 0) &= \pi \cdot (1 - \ell) + \pi \cdot \ell + (1 - \pi) \\ &= \pi(1 - \ell + \ell) + 1 - \pi \\ &= \pi(1) + 1 - \pi \\ &= 1 \ . \end{aligned}$$

We now compute the expeced value of the naive estimator of π to check whether the estimator is unbiased. We have

$$\begin{aligned} \mathbb{E}[\pi_{naive}] &= \mathbb{E}\left[\frac{\sum_{i=1}^n X_i}{n}\right] \\ &= \frac{1}{n} \cdot \sum_{i=1}^n \mathbb{E}[X_i] \\ &= \frac{1}{n} \cdot \sum_{i=1}^n (P(X_i = 1) \cdot 1 + P(X_i = 0) \cdot 0) \\ &= \frac{1}{n} \cdot \sum_{i=1}^n (P(X_i = 1)) \\ &= \frac{1}{n} \cdot \sum_{i=1}^n (\pi \cdot (1 - \ell)) \\ &= \frac{1}{n} \cdot n\,(\pi \cdot (1 - \ell)) \\ &= \pi \cdot (1 - \ell) \ . \end{aligned}$$

Therefore, the naive estimator of π is only unbiased whenever $\ell = 0$. Otherwise, the estimator is biased towards 0 the more the persons of the sensitive group lie ($\lim_{\ell \to 1} \mathbb{E}[\pi_{naive}] = 0$).

(h) *Derive expressions for the mean squared error (MSE) of the randomized response estimator and the naive estimator. Analyse the expressions to draw conclusions about when, depending on the relevant parameters* (π, a, ℓ, n)*, one estimator should be favored over the other.*

Problem 2

		$N = 500$	$N = 1000$	$N = 5000$
$\hat{N}$		590.45±311	1096.62±286	5152.14±583
Error (%)		18.09	9.66	3.04
$\hat{\sigma}_{\hat{N}}$		239.58±258	284.17±116	566.36±98
C.I.	$\alpha = 0.05$	[299.68±85; 1322.77±1298]	[687.34±137; 1842.05±609]	[4179.35±423; 6414.78±804]
	$\alpha = 0.01$	[250.36±59; 1740.30±2079]	[601.62±110; 2185.30±773]	[3922.46±383; 6884.52±890]

Table 1: Result from 1000 MC simulations for estimating $\hat{N}$. For each value, we report on the mean (black) and standard deviation (gray). We omit the digits after the decimal point for the standard deviation for brevity. The Error (%) is the absolute difference between the estimated and the true value of N, divided by the true value of N. Hence, it measures the relative error of the estimated value to the true value.

Observation. We implement the formulas and run the Monte-Carlo simulations for 1000 runs and report the results in Table 1. We first observe that, although the absolute error between the estimated and the true value of N increases with N, the relative error decreases when the true N is larger.

Interestingly, the given formula proposed for estimating the variance of $\hat{N}$ is very accurate, as its average value is very close to the computed variance of the Monte-Carlo simulation — in our case, we report on the standard deviation, which is simply the square root of the variance.

Lastly, we computed the averaged confidence intervals across all runs for $\alpha = 0.05$ and $\alpha = 0.01$. Naturally, the confidence intervals when using $\alpha = 0.01$ are larger than when using $\alpha = 0.05$, we notice that the standard deviation of the lower confidence decreases when increasing the size of the Confidence Interval (C.I.), whereas the standard deviation of the upper-confidence increases when decreasing α.

Discussion.

Statement on the Useage of Gen. AI

Unless stated otherwise, all code, text, and math derivations have been entirely thought and written by me with no external help — both for gen. AI and wolframalpha anc co.