\[
\frac{d}{dx} \sin^{-1} \sqrt{x} \quad \text{if} \quad \sin^{-1} \sqrt{x} = \theta \implies \sqrt{x} = \sin \theta \\
\implies x = \sin^2 \theta
\]

**Recap:**

Maximum likelihood estimator.

0 Bayes estimator  ② Minimax estimator

**Frequentist:**

\[x_1, \ldots, x_n \sim f_\theta(x)\]

Goal: provide point estimate of \(\theta\).

**Frequentist**

data are i.i.d.\(\text{r.v.'s}\), and parameter \(\theta\) is fixed but unknown.

**Bayesian**

data are fixed, parameter \(\theta\) is in a random variable.

Bayesian: Goal is to estimate the unknown distribution of \(\theta\).  

**Algorithm**:

① Start with a prior distribution on \(\theta\). This represents our prior belief on \(\theta\).

② Use this prior distribution and the data to find the posterior distribution of \(\theta\).

**Algorithm**:

Let \(x_1, \ldots, x_n \sim f_\theta(x)\)

Let the prior dist. of \(\theta\) be given by \(\pi(\theta)\).

The posterior dist. of \(\theta \mid x_1, \ldots, x_n\), denoted by \(\pi(\theta \mid x_1, \ldots, x_n)\) is defined as
\[ \pi(\theta | x_1, \ldots, x_n) = \frac{\left[ \prod_{i=1}^{n} f(x_i) \right] \pi(\theta)}{\int \left[ \prod_{i=1}^{n} f_0(x_i) \right] \pi(\theta) \, d\theta} \]

**Example:** \( x_1, \ldots, x_n \sim \text{Ber}(p) \).

**Goal:** Posterior dist. of \( p \).

Since \( \alpha p < 1 \), a reasonable prior dist. on \( p \) is given by \( p \sim \text{Beta}(\alpha, \beta) \).

\[
\pi(p) = \frac{p^{\alpha-1} (1-p)^{\beta-1}}{\text{Beta}(\alpha, \beta)}, \quad 0 < p < 1
\]

\[ \pi(p | x_1, \ldots, x_n) = \frac{\left[ \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} \right] \pi(p)}{\int \left[ \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} \right] \pi(p) \, dp} \]

\[ = \frac{p^{\sum_{i=1}^{n} x_i} (1-p)^n - \sum_{i=1}^{n} x_i} {\int p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i} p^{\alpha-1} (1-p)^{\beta-1} \, dp} \]

\[ = \frac{p^{\sum_{i=1}^{n} x_i - 1} (1-p)^{\beta + n - \sum_{i=1}^{n} x_i - 1}} {\int p^{\alpha + \sum_{i=1}^{n} x_i - 1} (1-p)^{\beta + n - \sum_{i=1}^{n} x_i - 1} \, dp} \]

\( \Rightarrow p | x_1, \ldots, x_n \sim \text{Beta}(\alpha + \sum_{i=1}^{n} x_i, \beta + n - \sum_{i=1}^{n} x_i) \)

\( 2 \)
Example: \( x_1, \ldots, x_n \sim \text{Pois}(\lambda), \quad \lambda \sim \text{Gamma}(a, b) \)

\[
\pi(\lambda) = \frac{\lambda^{a-1} e^{-\lambda b}}{\Gamma(a)}, \quad \lambda > 0
\]

\[
\pi(x_1, \ldots, x_n) = \frac{\left[ \prod_{i=1}^{n} \frac{e^{-\lambda x_i}}{x_i!} \right] \lambda^{a-1} e^{-\lambda b} b^a}{\Gamma(a)}
\]

\[
\frac{\int \left[ \prod_{i=1}^{n} \frac{e^{-\lambda x_i}}{x_i!} \right] \lambda^{a-1} e^{-\lambda b} b^a d\lambda}{\Gamma(a)}
\]

\[
e^{- (n+b) \lambda} \lambda^\frac{\sum x_i}{a} + a - 1
\]

\[
\int e^{- (n+b) \lambda} \lambda^\frac{\sum x_i}{a} + a - 1 d\lambda
\]

\[
\lambda | x_1, \ldots, x_n \sim \text{Gamma} \left( \frac{\sum x_i + a b}{a} \right), \quad n+b
\]

Conjugate family: Poisson

**Let** \( F \) **denote** the class of priors on \( \mu \). **A class** \( T \) **of prior distributions in a conjugate family for** \( F \) **if the posterior distribution is in the class** \( T \) **for all** \( f \in \mathcal{F} \), **all priors in** \( T \) **and all** \( x \in \mathcal{X} \).**

Now we will learn how to create a good estimator using this Bayesian knowledge.

**Let** \( \delta(x) \) **be an estimator of** \( \theta \). **The loss in estimating** \( \theta \) **by** \( \delta(x) \) **is represented by a function known as the loss function**. **We denote the loss fn. by** \( L(\theta, \delta) / L(\theta, \delta(x)) \)
\[
R(\theta, \delta) = \text{Ex}_{\theta} \left[ L(\theta, \delta(x)) \right]
\]

is known as the risk function of \( \delta \).

\[
L(\theta, \delta(x)) = (\theta - \delta(x))^2 \rightarrow \text{squared error loss function}
\]

\[
R(\theta, \delta) = \text{Ex}_{\theta} \left[ L(\theta, \delta(x)) \right] = \text{Ex}_{\theta} \left[ (\theta - \delta(x))^2 \right] \rightarrow \text{Mean squared error}
\]

Given two estimators \( \delta_1 \) and \( \delta_2 \), we say \( \delta_1 \) is better than \( \delta_2 \) if

\[
R(\theta, \delta_1) \leq R(\theta, \delta_2) \quad \forall \theta \quad \text{and the inequality is strict at least for one } \theta.
\]

We want to create a summary of the entire risk function.

1. Average risk = \( E_\theta \left[ R(\theta, \delta(x)) \right] = E_\theta \text{Ex}_{\theta} \left[ L(\theta, \delta(x)) \right] \)

2. Supremum risk = \( \sup_{\theta} R(\theta, \delta(x)) \)

If \( \delta(x) \) is an estimator of \( \theta \) that minimizes the average risk over all estimators, then \( \delta(x) \) is called the Bayes estimator of \( \theta \).
Average risk = \( E_\theta E_{x|\theta} \left[ L(\theta, s(x)) \right] \)
= \( E_{x, \theta} \left[ L(\theta, s(x)) \right] \)
= \( E_x E_{\theta|x} \left[ L(\theta, s(x)) \right] \)

If \( E_{\theta|x} \left[ L(\theta, s(x)) \right] \) \( \neq \) \( x \)
then that in going to be the minimizer of the average risk.

Finding the Bayes estimator is equivalent to finding an estimator \( s(x) \) that minimizes
\( E_{\theta|x} \left[ L(\theta, s(x)) \right] \)

If \( L(\theta, s(x)) = (\theta - s(x))^2 \)

arguing \( E_{\theta|x} \left[ (\theta - s(x))^2 \right] \)

\( E_{\theta|x} \left[ (\theta - \hat{s}(x))^2 \right] = E_{\theta|x} \left[ (\theta - E_{\theta|x}(\theta) + E_{\theta|x}(\theta) - \hat{s}(x))^2 \right] \)
= \( E_{\theta|x} \left[ (\theta - E_{\theta|x}(\theta))^2 \right] + 2 E_{\theta|x} \left[ (\theta - E_{\theta|x}(\theta)) (E_{\theta|x}(\theta) - \hat{s}(x)) \right] \)
= \( E_{\theta|x} \left[ (\theta - E_{\theta|x}(\theta))^2 \right] + 2 (E_{\theta|x}(\theta) - \hat{s}(x)) \left[ E_{\theta|x}(\theta) - E_{\theta|x}(\theta) \right] \)
\( \geq E_{\theta|x} \left[ (\theta - E_{\theta|x}(\theta))^2 \right] \)

Thus \( \hat{s}(x) = E_{\theta|x}(\theta) \) in the Bayes estimator in this case.
\arg \min_{\delta(x)} E_{\theta \mid x} (\theta - \delta(x))^\gamma \\
\Rightarrow \frac{d}{ds} E_{\theta \mid x} (\theta - \delta(x))^\gamma = E_{\theta \mid x} \left( \frac{d}{ds} (\theta - \delta)^\gamma \right) = 2E_{\theta \mid x} (\theta - \delta) = 0 \\
\Rightarrow \delta = E_{\theta \mid x} (\theta) \\
L(\theta, \delta(x)) = \omega(\theta) (\theta - \delta(x))^\gamma \rightarrow \text{weighted loss fn.} \\
\arg \min_{\delta(x)} E_{\theta \mid x} \left( \omega(\theta) (\theta - \delta(x))^\gamma \right) \\
\Rightarrow \frac{d}{ds} E_{\theta \mid x} (\omega(\theta) (\theta - \delta)^\gamma) = 0 \\
\Rightarrow 2E_{\theta \mid x} (\omega(\theta) (\theta - \delta)) = 0 \Rightarrow \delta = \frac{E_{\theta \mid x} (\theta \omega(\theta))}{E_{\theta \mid x} (\omega(\theta))} \\
\Rightarrow \delta(x) = \frac{E_{\theta \mid x} (\theta \omega(\theta))}{E_{\theta \mid x} (\omega(\theta))} \\

\text{Example: } x_1, \ldots, x_n \sim \text{Beta}(\theta), \quad \theta \sim \text{Beta}(\alpha, \beta) . \\
\prod_{i=1}^{n} \theta(x_1, \ldots, x_n) = \frac{\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} (1 - \theta)^{n - \sum_{i=1}^{n} x_i + \beta - 1}}{\text{Beta}(\sum_{i=1}^{n} x_i, n - \sum_{i=1}^{n} x_i + \beta)} \\
E[\theta \mid x_1, \ldots, x_n] = \int_{0}^{1} \theta \left( \frac{\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} (1 - \theta)^{n - \sum_{i=1}^{n} x_i + \beta - 1}}{\text{Beta}(\sum_{i=1}^{n} x_i, n - \sum_{i=1}^{n} x_i + \beta)} \right) d\theta \\
= \frac{\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} (1 - \theta)^{n - \sum_{i=1}^{n} x_i + \beta - 1}}{\text{Beta}(\sum_{i=1}^{n} x_i, n - \sum_{i=1}^{n} x_i + \beta) \text{Beta}(\alpha + \sum_{i=1}^{n} x_i + 1, n - \sum_{i=1}^{n} x_i + \beta)}
\[
= \frac{\prod (\alpha + \sum_{i=1}^{n} x_i + 1) \prod (\alpha+n+\beta)}{\prod (\alpha+1+n+\beta)}
\]
\[
\prod (\alpha + \sum_{i=1}^{n} x_i) \prod (\alpha+n+\beta)
\]
\[
= \frac{\sum_{i=1}^{n} x_i + \alpha}{\alpha+n+\beta} \quad \text{(Using the fact: } \prod (\alpha+1) = \alpha \prod (\alpha) \text{)}
\]
\[
= \frac{n}{\alpha+n+\beta} + \frac{(\alpha+\beta)}{(\alpha+\beta+n)} \cdot \frac{\alpha}{(\alpha+\beta)}
\]
\[
\frac{\bar{x}}{\alpha+n+\beta}
\]

Posterior mean is the convex combination of data and prior.

Since \(E[\bar{x}] = \theta \) implies the above estimator cannot be an unbiased estimator of \(\theta\).

Thus, No unbiased estimator \(\hat{s}(x)\) of \(\theta\) can be a Bayes estimator unless \(E_\theta E_{\theta'}(\theta-s(x)) = 0\).

\[E_{\theta'}[s(x)] = \theta \quad \text{as } s(x) \text{ in unbiased}
\]

and
\[E_{\theta|x}(\theta|\theta) = s(x) \quad \text{as } s(x) \text{ in Bayes estimator.}\]
\[ E_\theta \mathbf{E}_{x|\theta} [s(x) \theta] = E_\theta [\theta \mathbf{E}_{x|\theta} [s(x)] ] = E_\theta [\theta^\gamma] \quad \ldots \ldots \text{0} \]

Also, \[ E_\theta E_{x|\theta} [s(x) \theta] = E_\theta [s(x) E_{x|\theta} (\theta)] = E_\theta [s(x)^\gamma] \quad \ldots \ldots \text{2} \]

Also, \[ E_\theta \mathbf{E}_{x|\theta} [s(x) \theta] = E_\theta E_{x|\theta} [s(x) \theta] \]

\[ \Rightarrow \text{by 0 \& 2} \]

\[ E_\theta \mathbf{E}_{x|\theta} [s(x)^\gamma] = E_\theta \mathbf{E}_{x|\theta} [s(x) \theta] = E_\theta [\theta^\gamma] \quad \ldots \ldots \text{3} \]

Hence,

\[ E_{\theta, x} [(s(x) - \theta)^\gamma] = E_{\theta, x} [\theta^\gamma - 2s(x) \theta + s(x)^\gamma] \]

\[ = E_\theta E_{x|\theta} [s(x)^\gamma] + E_\theta [\theta^\gamma] - 2 E_\theta E_{x|\theta} [s(x) \theta] \]

\[ = 0 \quad (\text{by 3}) \]

\[ \Rightarrow \text{the above equation needs to be satisfied for } s(x) \text{ as both Bayes and unbiased estimator, } \]

\[ s(x) \text{ to be a Bayes and unbiased estimator, the above equation needs to be satisfied.} \]

Example: \( x_1, \ldots, x_n \sim iid N(\mu, \sigma^2) \), \( \sigma^2 \) is known, \( \bar{x} \) in the UMPUE and

\[ E_{x|X} [(\bar{x} - \mu)^\gamma] = E_\mu \left( \frac{\sigma^\gamma}{n} \right) = \frac{\sigma^\gamma}{n} \neq 0. \]

\[ \Rightarrow \bar{x} \text{ can't be a Bayes estimator.} \]
Minimax estimator minimizes \( \sup_{\theta} \mathbb{E}_{\theta} [L(\theta, s(x))] \). Minimax estimator is difficult to find in general. But there are certain cases in which Bayes estimator becomes a minimax estimator.

**Least favorable distribution**

A prior dist. \( \pi(\theta) \) on \( \Theta \) is known to be a least favorable prior if

\[
\mathbb{E}_{\theta} \mathbb{E}_{\pi} [L(\theta, s(x))] \geq \mathbb{E}_{\theta} \mathbb{E}_{\pi'} [L(\theta, s_{\pi'}(x))]
\]

for all prior dist. \( \pi' \) on \( \Theta \). Here \( s_{\pi} \) and \( s_{\pi'} \) are Bayes estimators w.r.t. priors \( \pi \) and \( \pi' \) respectively.
Recap:

Least favorable distribution.

A prior dist. $\pi(\theta)$ of $\theta$ in a least favorable prior if
$E_\theta \text{Ex}_1[ L(\theta, \delta_\pi(x)) ] \geq E_\theta \text{Ex}_1[ L(\theta, \delta_{\pi'}(x)) ]$ for all
prior dist. $\pi'$ on $\theta$, $\delta_\pi$ and $\delta_{\pi'}$ are the Bayes estimators
Corresponding to priors $\pi$ and $\pi'$.

Result: If $\pi(\theta)$ is a prior dist. for which
$\int \text{Ex}_1 [ L(\theta, \delta_\pi) ] \pi(\theta) \, d\theta = \sup_\theta \text{Ex}_1 [ L(\theta, \delta_\pi) ]$, where $\delta_\pi$
in the Bayes estimator under the prior $\pi(\theta)$,
then
a) $\delta_\pi$ is minimax
b) $\pi$ is least favorable.

Proof:

a) For any estimator $\delta(x)$,
$\sup_\theta \text{Ex}_1 [ L(\theta, \delta(x)) ] \geq E_\theta \text{Ex}_1 [ L(\theta, \delta(x)) ]$
$\geq E_\theta \text{Ex}_1 [ L(\theta, \delta_\pi) ] = \sup_\theta \text{Ex}_1 [ L(\theta, \delta_\pi) ]$

Thus $\delta_\pi$ minimizes supremum risk over all estimators. Hence $\delta_\pi$ is minimax estimator.

b) Note that
$E_\theta \text{Ex}_1 [ L(\theta, \delta_{\pi'}(x)) ] = \sup_\theta \text{Ex}_1 [ L(\theta, \delta_{\pi'}) ]$
$\geq \int \text{Ex}_1 [ L(\theta, \delta_{\pi'}) ] \pi'(\theta) \, d\theta$ (For any other prior)
$\geq \int \text{Ex}_1 [ L(\theta, \delta_\pi) ] \pi'(\theta) \, d\theta$
$= E_\theta \text{Ex}_1 [ L(\theta, \delta_\pi) ]$
\[ \Rightarrow \text{\( p \) in the least favourable dist.} \]

**Example:** \( x_1, \ldots, x_n \overset{iid}{\sim} \text{Ber}(p), p \sim \text{Beta}(\alpha, \beta) \).

Bayes estimator of \( p \) under squared error loss

\[ \hat{p}_n(x) = \frac{\alpha + \sum_{i=1}^{n} x_i}{n + \alpha + \beta} \]

\[ E_x[p \left( (\hat{p}(x) - p)^2 \right)] = \frac{1}{(\alpha + \beta + n)} \left[ \alpha^2 + \left\{ n - 2\alpha(\alpha + \beta) \right\} + \left\{ (\alpha + \beta)^2 - n^2 \beta \right\} \right] \]

We want to make this risk fn. constant fn. of \( p \)

Thus we set,

\[ n = 2\alpha(\alpha + \beta) \quad \text{and} \quad (\alpha + \beta)^2 = n \]

\[ \Rightarrow \alpha = \frac{\sqrt{n}}{2}, \quad \beta = \frac{\sqrt{n}}{2} \]

Thus, under the prior dist., \( p \sim \text{Beta}(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}) \), Bayes estimator produces constant risk.

Now use the result to argue that this Bayes estimator \( \frac{\sqrt{n}}{2} + \sum_{i=1}^{n} x_i \) is a minimax estimator under the prior \( \text{Beta}(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}) \).

Of course, by the same result \( \text{Beta}(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}) \) is the least favourable dist.

3. **Testing of Hypothesis**

   **Statistical hypothesis testing is all about**

   1. **Beginning with a tentative idea about the unknown parameter.**

   2. Want to test validity of this tentative idea based on sample information.
3) Existing tentative idea: \( H_0 \) (null hypothesis), new idea: \( H_1 \) (alternative hypothesis).

4) We begin by assuming that the null hypothesis is true. Only when there is an overwhelming evidence contradicting the null do we reject it in favor of alternative.

<table>
<thead>
<tr>
<th>Do not reject ( H_0 )</th>
<th>( H_0 ) is true</th>
<th>( H_0 ) is false</th>
</tr>
</thead>
<tbody>
<tr>
<td>Censored</td>
<td></td>
<td>Type 2 error</td>
</tr>
<tr>
<td>reject ( H_0 )</td>
<td>Type 1 error</td>
<td></td>
</tr>
</tbody>
</table>

Our goal is to minimize
\[
P(\text{Type 1 error}) = P(\text{reject } H_0 \mid H_0 \text{ is true})
\]
\[
P(\text{Type 2 error}) = P(\text{not rejecting } H_0 \mid H_0 \text{ is false})
\]

They can't be minimized together. Thus we fix \( P(\text{type 1 error}) \) and minimize \( P(\text{Type 2 error}) \).

Minimizing \( P(\text{Type 2 error}) \)
\[
\Leftrightarrow \text{Maximizing} \quad 1 - P(\text{Type 2 error}) = P(\text{rejecting } H_0 \mid H_0 \text{ is false})
\]

\( P(\text{type 1 error}) \) is called the level of the test and \( 1 - P(\text{type 2 error}) \) is called the power of the test.
Parametric tests: Let $x_1, \ldots, x_n \sim f(x|\theta)$.

We test $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_1$, where $\Theta_0$ and $\Theta_1$ are disjoint sets. If $\Theta_0$ is a singleton set, then the null hypothesis is called a simple null hypothesis. If $\Theta_0$ has multiple elements, then the null hypothesis is called a composite null hypothesis.

Rejection region: Let $R = \{x \in X | H_0 \text{ is rejected for } x \}$ be known as the rejection region or critical region of a test.

Let $\phi(x) = \text{Prob. of rejecting } H_0 \text{ when } x \text{ is observed.}$

The power function of a test is given by

$\beta(\theta) = \mathbb{E}_{X|\theta} \phi(x)$, power function is a function of the parameter $\theta$.

Consider the situation $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$.

$\beta(\theta_0) = \mathbb{E}_{X|\theta_0} \phi(x)$ level of the test.

$\beta(\theta_1) = \mathbb{E}_{X|\theta_1} \phi(x)$ power of the test.

Under the composite hypothesis, $H_0: \theta \in \Theta_0$.

The level of the test $\alpha$ implies

$\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha.$
Given a certain level, we want to find the most powerful test.

<table>
<thead>
<tr>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0$</td>
<td>$1/8$</td>
<td>$1/8$</td>
<td>$1/4$</td>
<td>$1/2$</td>
</tr>
<tr>
<td>$f_1$</td>
<td>$1/2$</td>
<td>$1/4$</td>
<td>$1/8$</td>
<td>$1/8$</td>
</tr>
</tbody>
</table>

$\forall f_0$, $\theta = 0, 1$

- How to find the best (most powerful level $1/8$) test.
- Level $1/8$ means rejection region $R$ must have prob. $1/8$ under $H_0$.

$R_1 = \{0, 2\}$, test 1.

$R_2 = \{1\}$, test 2.

Power of test 1: $P(\text{rejecting } H_0 \mid H_0 \text{ is false}) = P(X \in R_1 \mid H_0 \text{ is false}) = 1/2$.

Power of test 2: $P(X \in R_2 \mid H_0 \text{ is false}) = 1/4$.

Test 1 is more powerful than test 2, under the same level.

$f_1$ is higher at 0 than at 1.

Let us decide to reject $H_0$ with $f_0$ a rejection region that contains high values of $f_1$.

$\text{Power} = P(X \in R \mid f_1 \text{ is true})$
Theorem: Neyman-Pearson Lemma:

Consider testing \( H_0 : \theta = \theta_0 \) vs. \( H_1 : \theta = \theta_1 \), where pdf or pmf corresponding to \( \theta_i \) is \( f(x|\theta_i) \), \( i=0,1 \).

Using a test with rejection region \( R \) that satisfies

\[
\phi(x) = \begin{cases} 
1 & \text{if } \frac{f(x|\theta_1)}{f(x|\theta_0)} > k \\
0 & \text{otherwise}
\end{cases}
\]

for some \( k > 0 \), and \( \alpha = P_{\theta_0}(x \in R) \). Then

a) Any test that satisfies the above is the most powerful level \( \alpha \)-test.

b) If there exists a test satisfying the above, then it is a \( \theta \) most powerful level \( \alpha \)-test of \( \theta_0 \) level.

Note:

\[
\frac{f(x|\theta_1)}{f(x|\theta_0)} = \frac{g(T(x)|\theta_0) h(x)}{g(T(x)|\theta_0) h(x)}
\]

MP test can be written as

\[
\phi(x) = \begin{cases} 
1 & \text{if } g(T(x)|\theta_1) > k g(T(x)|\theta_0) \\
0 & \text{otherwise}
\end{cases}
\]

Example: \( X_1, X_2 \sim \text{iid Bern}(\theta) \). Want to test \( H_0 : \theta = \frac{1}{2} \) vs. \( H_1 : \theta = \frac{3}{4} \).

\[ \sum_{i=1}^{2} X_i = \text{sufficient Stat.} \]

\[ \sum_{i=1}^{2} X_i \sim \text{Bin}(2, \theta) \]
\[ \frac{f(0 \mid \theta = \frac{3}{4})}{f(0 \mid \theta = \frac{1}{2})} = \frac{1}{4} , \quad \frac{f(1 \mid \theta = \frac{3}{4})}{f(1 \mid \theta = \frac{1}{2})} = \frac{3}{4} , \quad \frac{f(2 \mid \theta = \frac{3}{4})}{f(2 \mid \theta = \frac{1}{2})} = \frac{9}{4} . \]

If we choose \( \frac{3}{4} < k < \frac{9}{4} \) \( \Rightarrow R = \{ 2 \} \)

MP level a-test for \( a = P \left( \sum_{i=1}^{k} x_i = 2 \mid \theta = \frac{1}{2} \right) = \frac{1}{4} . \)