Capture Recapture estimation using finite mixtures of arbitrary dimension

Richard Arnold
Yu Hayakawa
Paul Yip

International Biometrics Society
Taupo, 30 Nov-3 Dec 2009

Outline

- Capture-Recapture experiments and models
- Reversible Jump MCMC
- Application: Software reliability

Capture-Recapture

- k repeated samples taken from a population
- Population size N is unknown
- D distinct individuals are seen
- Some seen on multiple occasions, some only once
- **Goal:** estimate population size N

Capture-Recapture: Applications

- **Size of animal populations**:
  Samples are occasions on which animals are trapped, marked (for re-identification) and released.
  N is the number of animals in the population.

Capture-Recapture: Applications

- **Size of animal populations**:
  Samples are occasions on which animals are trapped, marked (for re-identification) and released.
  N is the number of animals in the population.

- **Software testing**:
  Samples are independent software testers detecting errors.
  N is the number of errors in the piece of software.

Observations: Capture Matrix X

	Samples					
		1	2	3	...	k
Individuals	1	1	0	1	...	0
	2	1	0	1	...	0
	3	0	0	0	...	0
Observed	4	0	1	1	...	0
	5	0	0	1	...	0
	$\vdots$					
	D	0	1	1	...	0
	$D+1$	0	0	0	...	0
	$D+2$	0	0	0	...	0
Unobserved	$\vdots$					
	N	0	0	0	...	0

An unknown number $(N-D)$ of zero rows: to be estimated.

Capture-Recapture

Observations form the $D \times k$ **capture matrix**

$$X_{ij} = \begin{cases} 1 & \text{if individual } i \text{ appears in sample } j \\ 0 & \text{otherwise} \end{cases}$$

Probability of capture

$$\begin{aligned} p_{ij} &= \text{Probability individual } i \text{ is captured in sample } j \\ &= p(X_{ij} = 1) \end{aligned}$$

$$X_{ij} \sim \text{Bernoulli}(p_{ij})$$

$$p(X|P,N) = \frac{N!}{\prod_{\mathbf{x}} N_{\mathbf{x}}!} \prod_{i=1}^{N} \prod_{j=1}^{k} p_{ij}^{x_{ij}} (1-p_{ij})^{1-x_{ij}}$$

Some key assumptions

1. All individuals are catchable
2. **Closed population**: no births/deaths or migration
3. No loss of marks (identifiable individuals)
4. No impact of sampling on capture probabilities
(i.e. no behavioural responses)

(Models do exist for violations of 2-4)

What makes this a difficult problem?

What makes this a difficult problem?

1. Heterogeneity

What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model M_{h}

What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model M_{h}
- **Samples differ** – some samples are more successful: model M_{t}

What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model M_{h}
- **Samples differ** – some samples are more successful: model M_{t}
- **Both sources** – model M_{th}

What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model M_{h}
- **Samples differ** – some samples are more successful: model M_{t}
- **Both sources** – model M_{th}

2. Identifiability

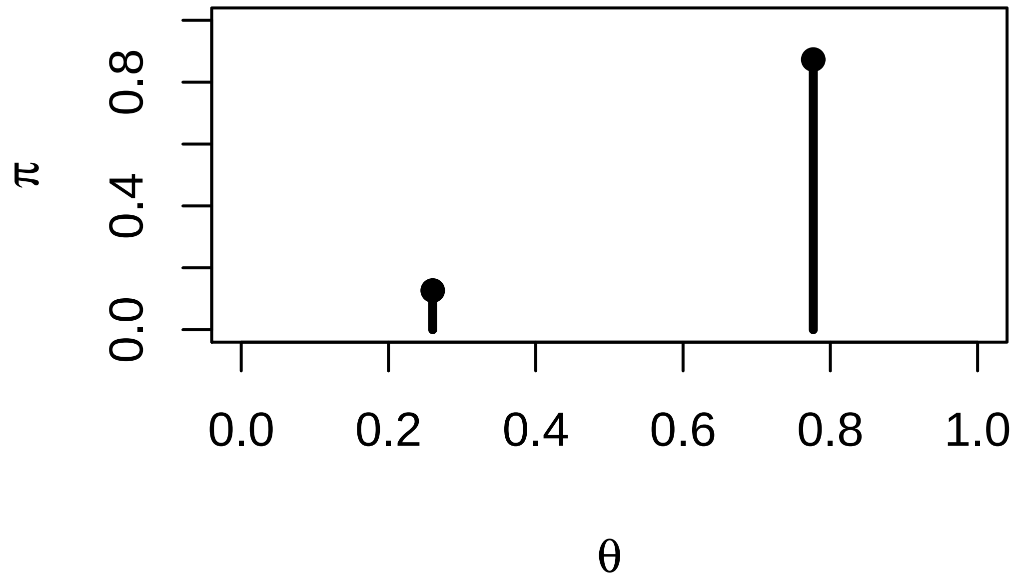
- N is identifiable **within** but not **between** model classes.
  e.g. finite vs. beta (infinite) mixture models for P

Modelling Heterogeneity by Finite Mixtures

1. **Individuals** – animals/errors differ in their catchability

Individuals belong to A latent classes:

- Membership probabilities $\{\pi_a\}$
- Capture probabilities $\{\phi_a\}$



$$f(p_i) = \sum_{a=1}^{A} \pi_a \delta(p_i - \phi_a) \qquad \sum_{a=1}^{A} \pi_a = 1$$

Modelling Heterogeneity by Finite Mixtures

2. **Samples** – probability of capture varies between samples
   - **Model** $M_{\mathrm{t[f]h}}$: k fixed effects

$$p_{ij} = \theta_{aj}$$
$$\mathrm{logit}(\theta_{aj}) = \mathrm{logit}(\phi_a) + \beta_j + \gamma_{aj}$$

Individual i is in latent class a

Modelling Heterogeneity by Finite Mixtures

2. **Samples** – probability of capture varies between samples
   - **Model** $M_{\mathrm{t[f]h}}$: k fixed effects

$$p_{ij} = \theta_{aj}$$
$$\mathrm{logit}(\theta_{aj}) = \mathrm{logit}(\phi_a) + \beta_j + \gamma_{aj}$$

Individual i is in latent class a

OR

Modelling Heterogeneity by Finite Mixtures

2. **Samples** – probability of capture varies between samples
   - **Model** $M_{\mathrm{t[f]h}}$: k fixed effects

$$p_{ij} = \theta_{aj}$$
$$\mathrm{logit}(\theta_{aj}) = \mathrm{logit}(\phi_a) + \beta_j + \gamma_{aj}$$

Individual i is in latent class a

OR

   - **Model** M_{th} B latent classes: membership probabilities $\{\lambda_b\}$ capture probabilities $\{\psi_b\}$

$$p_{ij} = \theta_{ab}$$
$$\mathrm{logit}(\theta_{ab}) = \mathrm{logit}(\phi_a) + \mathrm{logit}(\psi_b) - \mathrm{logit}(\phi_1) + \gamma_{ab}$$

Individual i is in latent class a and
Sample j is in latent class b

Identifiability

- Ordering constraints on mixture component probabilities

$$0 < \phi_1 < \ldots < \phi_A < 1 \;\text{ and }\; \phi_1 = \psi_1 < \psi_2 < \ldots < \psi_B < 1$$

  Implemented by 'repulsive' priors (following Green 1995)

$$\underline{\phi}|A \sim \mathrm{EOS}(\mathrm{Uniform}(0,1)^{2A+1})$$

  $\{\phi_a\}$ are the even order statistics of $(2A+1)$ draws from $U(0,1)$
- Sum to zero constraints on fixed effects

$$\begin{aligned} \sum_j \beta_j &= \sum_a \gamma_{aj} = \sum_j \gamma_{aj} = 0 \quad \text{or} \\ \sum_a \gamma_{ab} &= \sum_b \gamma_{ab} = 0 \end{aligned}$$

  Ensured by degenerate Normal priors, e.g.

$$\beta_j \sim \mathrm{Normal}(0, {\sigma'_\beta}^2) \quad \text{with} \sum_j \beta_j = 0$$

Priors

			Models	
Parameter	Distribution	Constants	$M_{\text{t[f]h}}$	M_{th}
N	Geometric$(1-\eta)$	$\eta = 0.999$	$\star$	$\star$
A	Geometric$(1-\rho_h)$	$\rho_h = 0.8$	$\star$	$\star$
B	Geometric$(1-\rho_h)$	$\rho_h = 0.8$		$\star$
$\underline{\pi}\|A$	Dirichlet$(\underline{1}\alpha_h)$	$\alpha_h = \frac{3}{2}$	$\star$	$\star$
$\underline{\phi}\|A$	EOS$(\text{Uniform}(0,1)^{2A+1})$		$\star$	$\star$
σ^2_β	InverseGamma$(\nu_\beta, \kappa_\beta)$	$\nu_\beta = 3, \kappa_\beta = 40$	$\star$	
$\underline{\beta}$	DegenNormal$(k;0,\sigma^2_\beta)$		$\star$	
$\underline{\lambda}\|B$	Dirichlet$(\underline{1}\alpha_t)$	$\alpha_t = \frac{3}{2}$		$\star$
$\underline{\psi}\|B,\phi_1$	EOS$(\text{Uniform}(\phi_1,1)^{2B-1})$			$\star$
$\zeta\|A,(B)$	Bernoulli(p_γ)	$p_\gamma = \frac{1}{2}$	$\star$	$\star$
$\gamma\|\zeta,A,(B)$	DegenNormal$(A,B;0,\sigma^2_\gamma)$	$\sigma_\gamma = 5$	$\star$	$\star$
$\underline{b}\|B,\underline{\lambda}$	Discrete$(\{1,\ldots,B\};\underline{\lambda})$			$\star$

Bayesian Estimation – RJMCMC

Estimation by Reversible Jump MCMC

- Draw from posterior
$$p(\{m, \Theta_m\}|X)$$
- Reversible Jump MCMC allows dimension switching moves: Addition/deletion of mixture components.
- 4 types of MCMC move in model M_{h}.

RJMCMC Updates

MCMC Step	1				
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$				

RJMCMC Updates

MCMC Step	1	2			
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$			

RJMCMC Updates

MCMC Step	1	2	3		
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$		

RJMCMC Updates

MCMC Step	1	2	3	4	
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$	$\theta_1^{(4)}$ $\theta_2^{(4)}$ $\theta_3^{(4)}$	

RJMCMC Updates

MCMC Step	1	2	3	4	…
	$\theta_1^{(1)}$	$\theta_1^{(2)}$	$\theta_1^{(3)}$	$\theta_1^{(4)}$	…
	$\theta_2^{(1)}$	$\theta_2^{(2)}$	$\theta_2^{(3)}$	$\theta_2^{(4)}$	…
	$\theta_3^{(1)}$	$\theta_3^{(2)}$	$\theta_3^{(3)}$	$\theta_3^{(4)}$	…

RJMCMC Updates

MCMC Step	1	2	3	4	...
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$	$\theta_1^{(4)}$ $\theta_2^{(4)}$ $\theta_3^{(4)}$	... ...

RJMCMC Step	1				
	$m^{(1)}$ $\theta_{m^{(1)}:1}^{(1)}$ $\theta_{m^{(1)}:2}^{(1)}$				

RJMCMC Updates

MCMC Step	1	2	3	4	…
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$	$\theta_1^{(4)}$ $\theta_2^{(4)}$ $\theta_3^{(4)}$	… … …

RJMCMC Step	1	2			
	$m^{(1)}$ $\theta^{(1)}_{m^{(1)}:1}$ $\theta^{(1)}_{m^{(1)}:2}$	$m^{(2)}$ $\theta^{(2)}_{m^{(2)}:1}$ $\theta^{(2)}_{m^{(2)}:2}$ $\theta^{(2)}_{m^{(2)}:3}$ $\theta^{(2)}_{m^{(2)}:4}$			

RJMCMC Updates

MCMC Step	1	2	3	4	...
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$	$\theta_1^{(4)}$ $\theta_2^{(4)}$ $\theta_3^{(4)}$	... ...

RJMCMC Step	1	2	3		
	$m^{(1)}$ $\theta^{(1)}_{m^{(1)}:1}$ $\theta^{(1)}_{m^{(1)}:2}$	$m^{(2)}$ $\theta^{(2)}_{m^{(2)}:1}$ $\theta^{(2)}_{m^{(2)}:2}$ $\theta^{(2)}_{m^{(2)}:3}$ $\theta^{(2)}_{m^{(2)}:4}$	$m^{(3)}$ $\theta^{(3)}_{m^{(3)}:1}$ $\theta^{(3)}_{m^{(3)}:2}$ $\theta^{(3)}_{m^{(3)}:3}$ $\theta^{(3)}_{m^{(3)}:4}$		

RJMCMC Updates

MCMC Step	1	2	3	4	...
	$\theta_1^{(1)}$ $\theta_2^{(1)}$ $\theta_3^{(1)}$	$\theta_1^{(2)}$ $\theta_2^{(2)}$ $\theta_3^{(2)}$	$\theta_1^{(3)}$ $\theta_2^{(3)}$ $\theta_3^{(3)}$	$\theta_1^{(4)}$ $\theta_2^{(4)}$ $\theta_3^{(4)}$	... ...

RJMCMC Step	1	2	3	4	
	$m^{(1)}$ $\theta^{(1)}_{m^{(1)}:1}$ $\theta^{(1)}_{m^{(1)}:2}$	$m^{(2)}$ $\theta^{(2)}_{m^{(2)}:1}$ $\theta^{(2)}_{m^{(2)}:2}$ $\theta^{(2)}_{m^{(2)}:3}$ $\theta^{(2)}_{m^{(2)}:4}$	$m^{(3)}$ $\theta^{(3)}_{m^{(3)}:1}$ $\theta^{(3)}_{m^{(3)}:2}$ $\theta^{(3)}_{m^{(3)}:3}$ $\theta^{(3)}_{m^{(3)}:4}$	$m^{(4)}$ $\theta^{(4)}_{m^{(4)}:1}$ $\theta^{(4)}_{m^{(4)}:2}$	

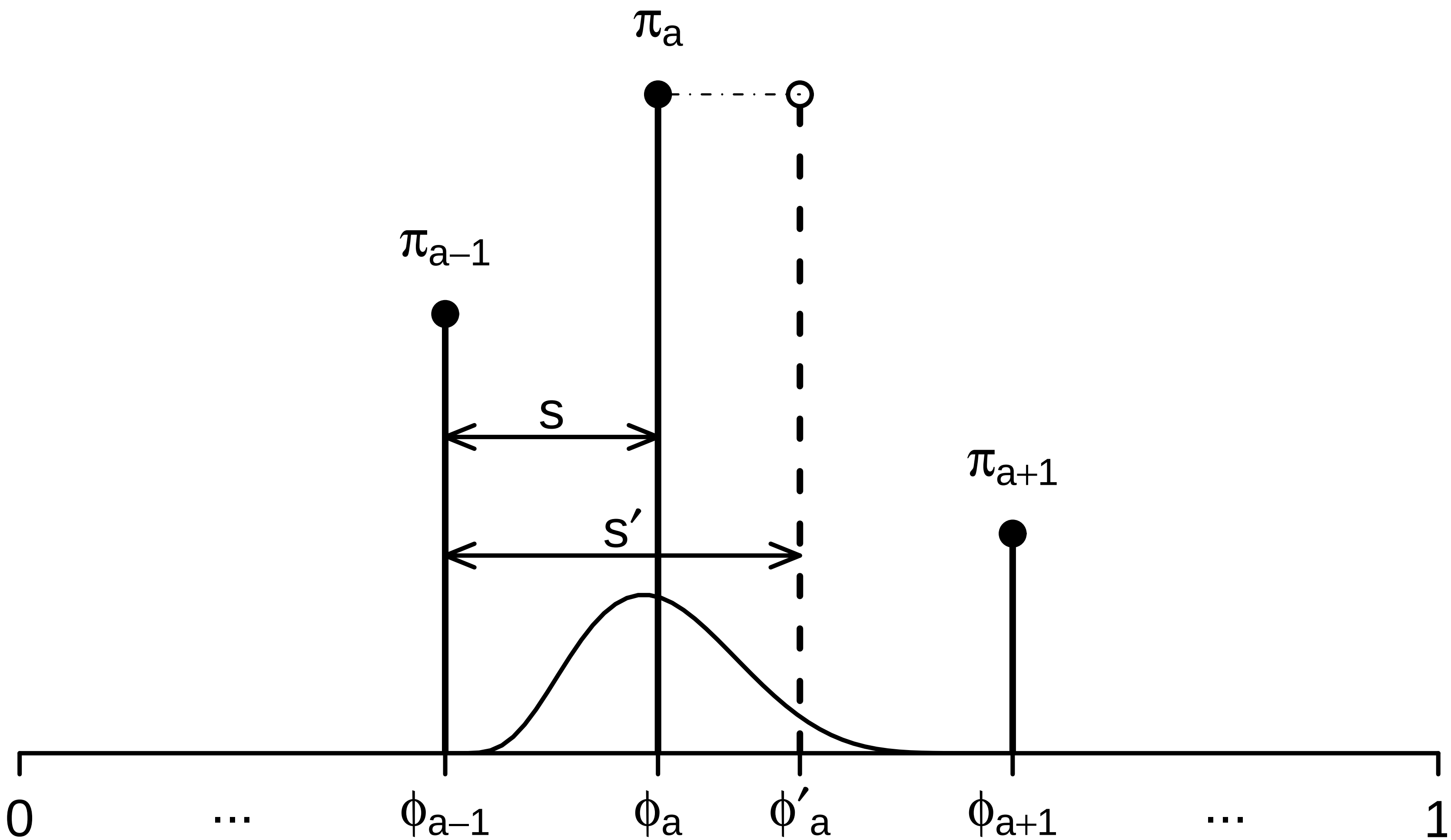
RJMCMC Updates

MCMC Step	1	2	3	4	...
	$\theta_1^{(1)}$	$\theta_1^{(2)}$	$\theta_1^{(3)}$	$\theta_1^{(4)}$	...
	$\theta_2^{(1)}$	$\theta_2^{(2)}$	$\theta_2^{(3)}$	$\theta_2^{(4)}$	...
	$\theta_3^{(1)}$	$\theta_3^{(2)}$	$\theta_3^{(3)}$	$\theta_3^{(4)}$	...

RJMCMC Step	1	2	3	4	...
	$m^{(1)}$	$m^{(2)}$	$m^{(3)}$	$m^{(4)}$	...
	$\theta^{(1)}_{m^{(1)}:1}$	$\theta^{(2)}_{m^{(2)}:1}$	$\theta^{(3)}_{m^{(3)}:1}$	$\theta^{(4)}_{m^{(4)}:1}$	...
	$\theta^{(1)}_{m^{(1)}:2}$	$\theta^{(2)}_{m^{(2)}:2}$	$\theta^{(3)}_{m^{(3)}:2}$	$\theta^{(4)}_{m^{(4)}:2}$	...
		$\theta^{(2)}_{m^{(2)}:3}$	$\theta^{(3)}_{m^{(3)}:3}$		...
		$\theta^{(2)}_{m^{(2)}:4}$	$\theta^{(3)}_{m^{(3)}:4}$		...

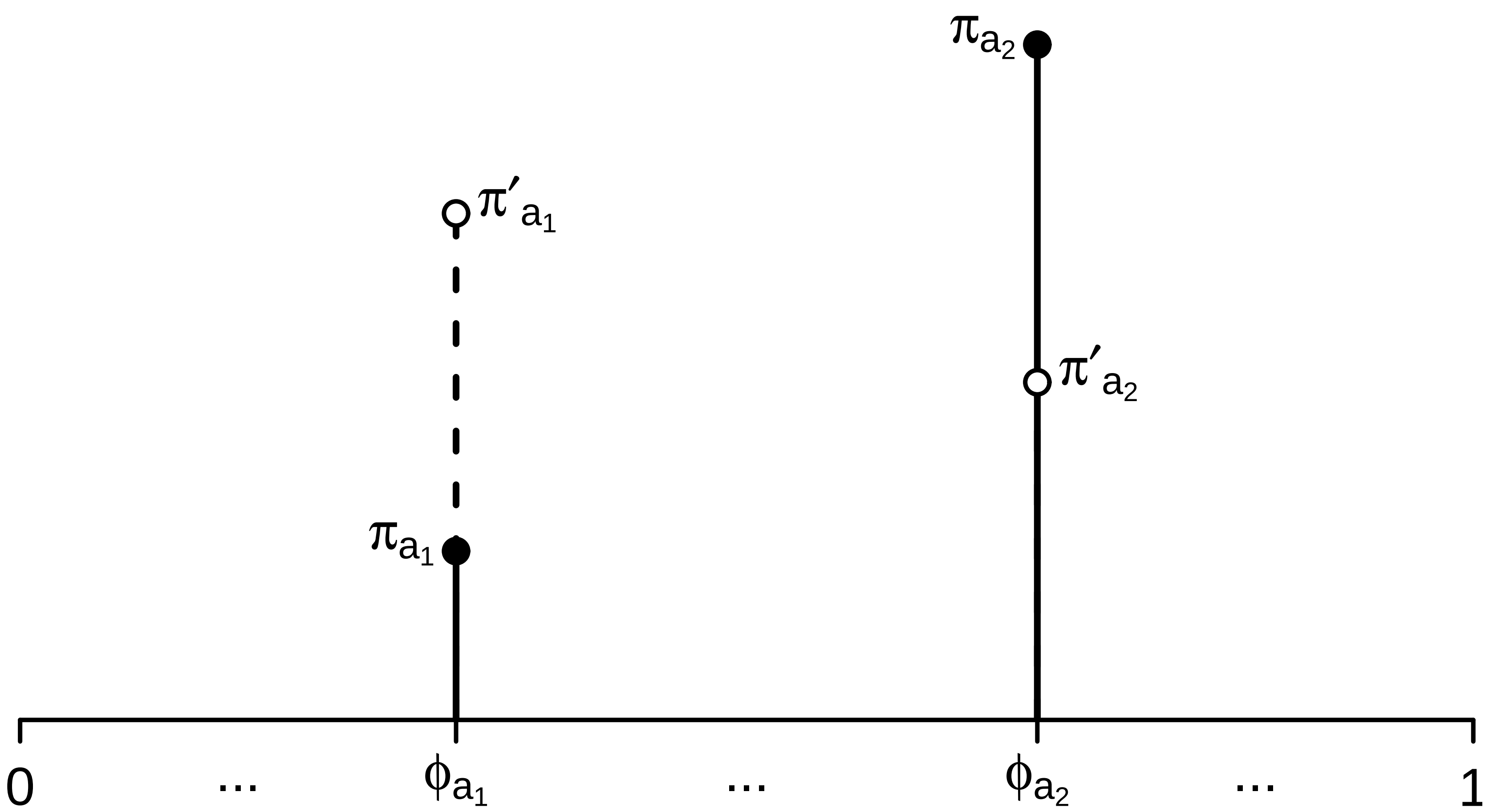
RJMCMC Updates

(1) Shift a support point



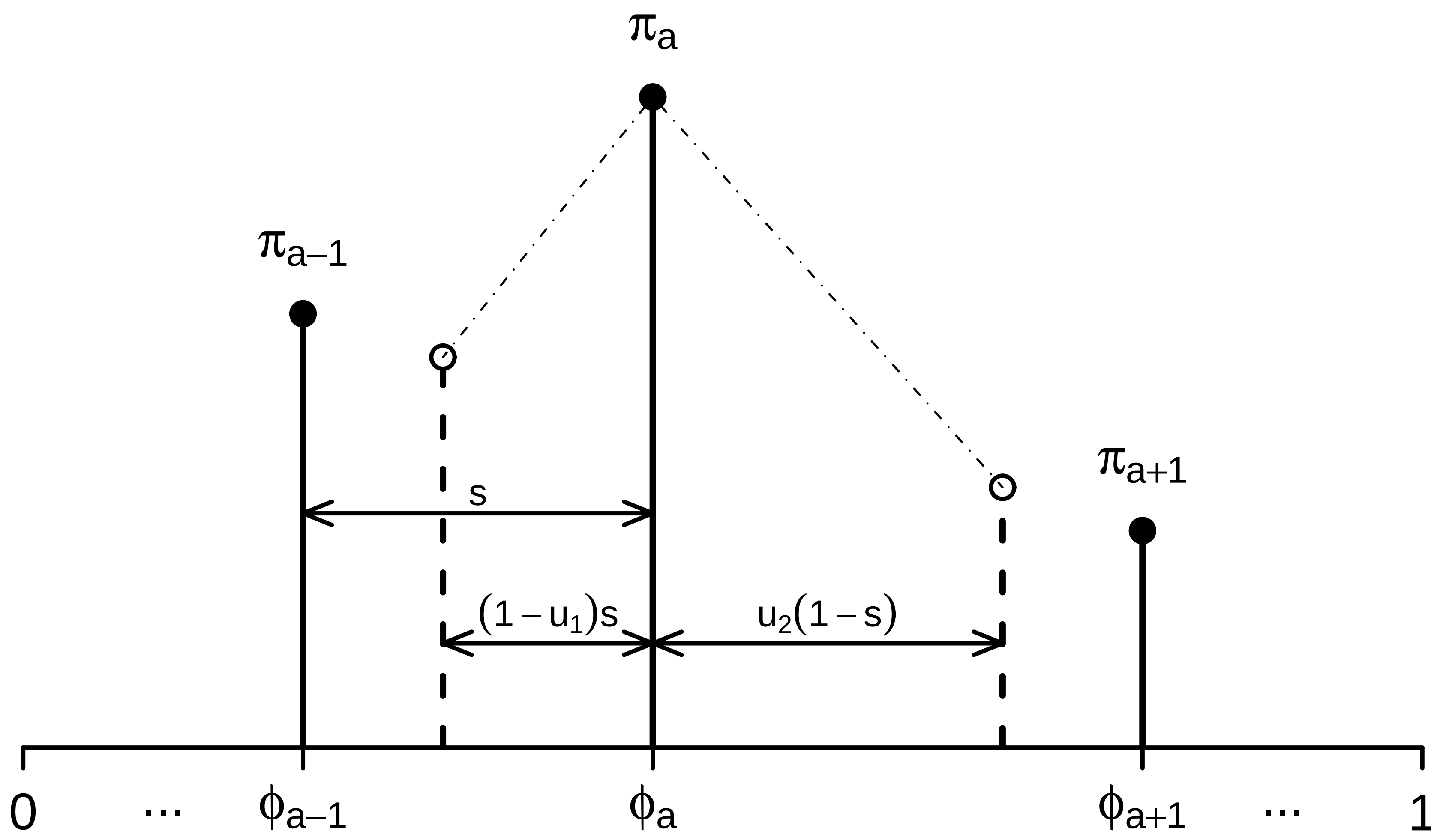
RJMCMC Updates

(2) Exchange probability between points



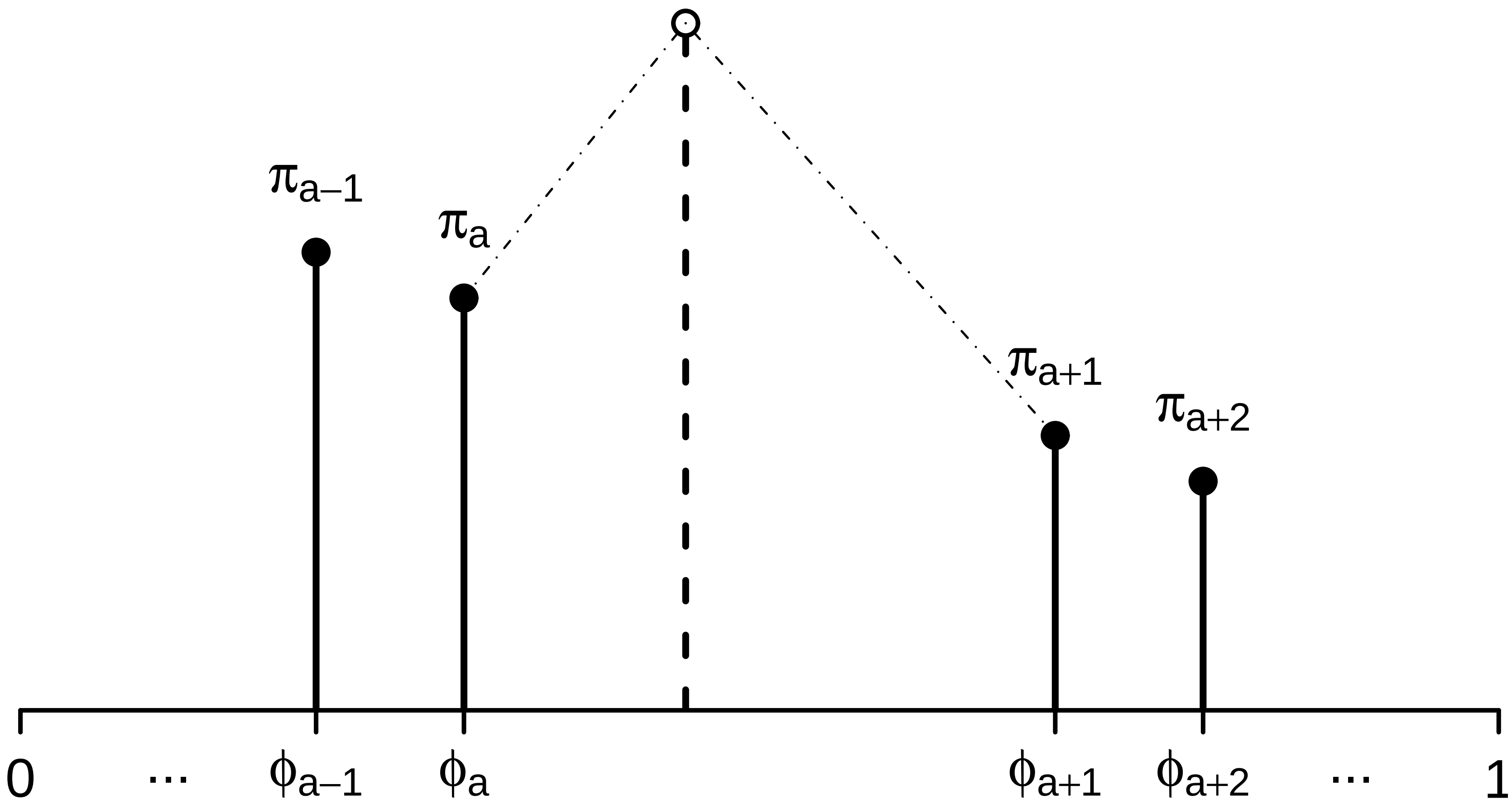
RJMCMC Updates

(3) Split a support point



RJMCMC Updates

(4) Merge two support points



Label switching and poor mixing

Estimation of finite mixture models affected if the MCMC sampler **can't mix** because of

- artificial identifiability constraints,
  (e.g. so that individuals/samples become persistently misallocated) or
- updating protocols making one mode inaccessible from another.

Label switching and poor mixing

Estimation of finite mixture models affected if the MCMC sampler **can't mix** because of

- artificial identifiability constraints,
  (e.g. so that individuals/samples become persistently misallocated) or
- updating protocols making one mode inaccessible from another.

In our case:

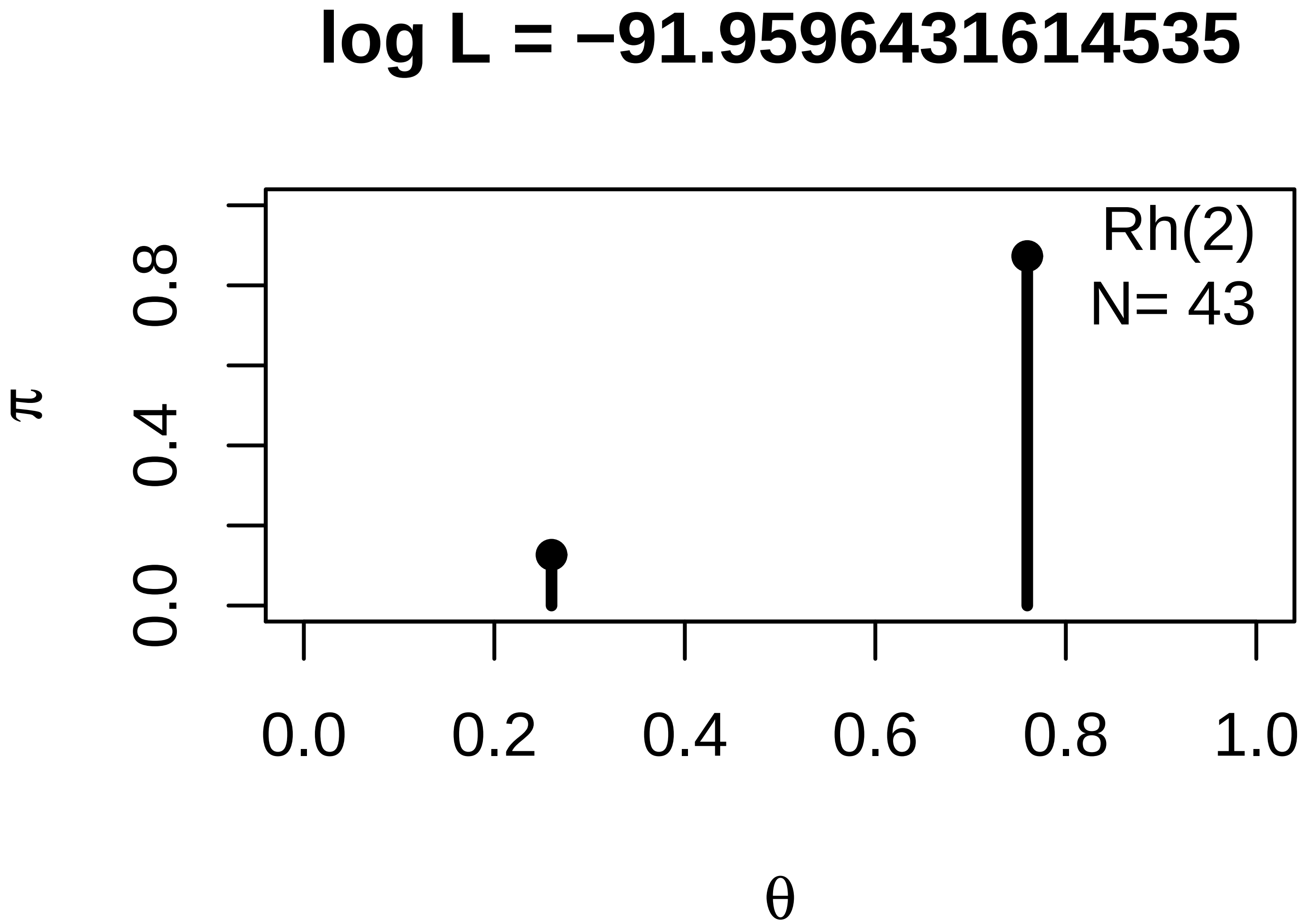
- Mixture components are labelled by a **single parameter**:
  have a natural ordering (e.g. preventing misallocations);
- Dimension-switching RJMCMC allows components to be deleted and reappear – the chain moves rapidly around model space.

Application: AT&T Switch Testing

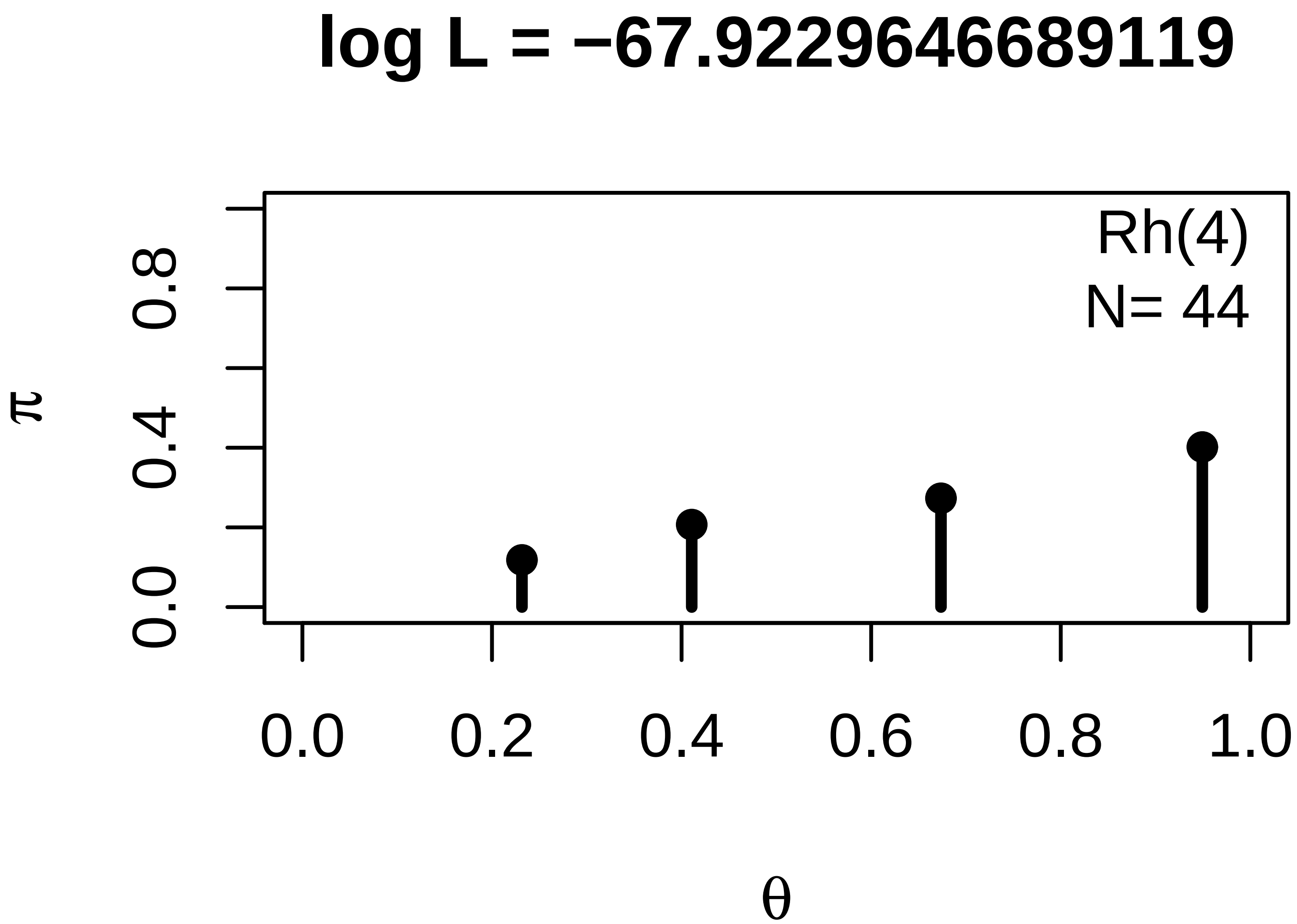
$k = 6$ reviewers tested switches and found $D = 43$ errors (Basu 1998)

Error	Reviewer j						
i	1	2	3	4	5	6	x_i
1	1	0	0	0	0	0	1
2	0	0	0	1	0	1	2
3	0	0	0	0	1	0	1
4	1	0	0	0	0	0	1
5	0	0	0	1	0	0	1
6	0	0	0	1	0	0	1
7	0	0	0	1	0	0	1
8	1	0	0	0	0	0	1
9	1	0	0	0	0	0	1
10	0	0	0	1	0	0	1
11	1	0	0	1	0	0	2
12	1	0	0	0	0	0	1
13	1	0	0	1	0	0	2
14	1	0	0	0	0	0	1
15	1	0	0	1	0	0	2
16	1	0	0	1	0	0	2
17	1	1	0	1	1	1	5
18	1	0	0	0	0	1	2
19	0	1	0	0	0	0	1
20	0	1	0	0	0	0	1
21	1	0	0	1	0	0	2
22	1	0	0	1	0	0	2
23	0	0	0	0	0	1	1
24	0	0	0	0	1	0	1
25	0	0	0	1	1	0	2
26	1	0	0	0	0	0	1
27	0	0	0	0	1	0	1
28	0	0	1	0	0	0	1
29	1	0	0	0	0	0	1
30	1	0	0	1	1	0	3
31	1	0	0	1	0	0	2
32	1	0	0	0	0	0	1
33	1	0	0	0	0	0	1
34	0	0	1	0	0	0	1
35	0	0	1	0	0	0	1
36	0	0	0	0	0	1	1
37	0	0	1	0	0	0	1
38	1	0	0	0	0	1	2
39	1	0	0	0	0	0	1
40	1	0	0	0	0	0	1
41	0	0	0	0	1	0	1
42	1	0	0	0	0	0	1
43	1	0	0	0	0	0	1
n_j	25	3	4	15	7	6	60

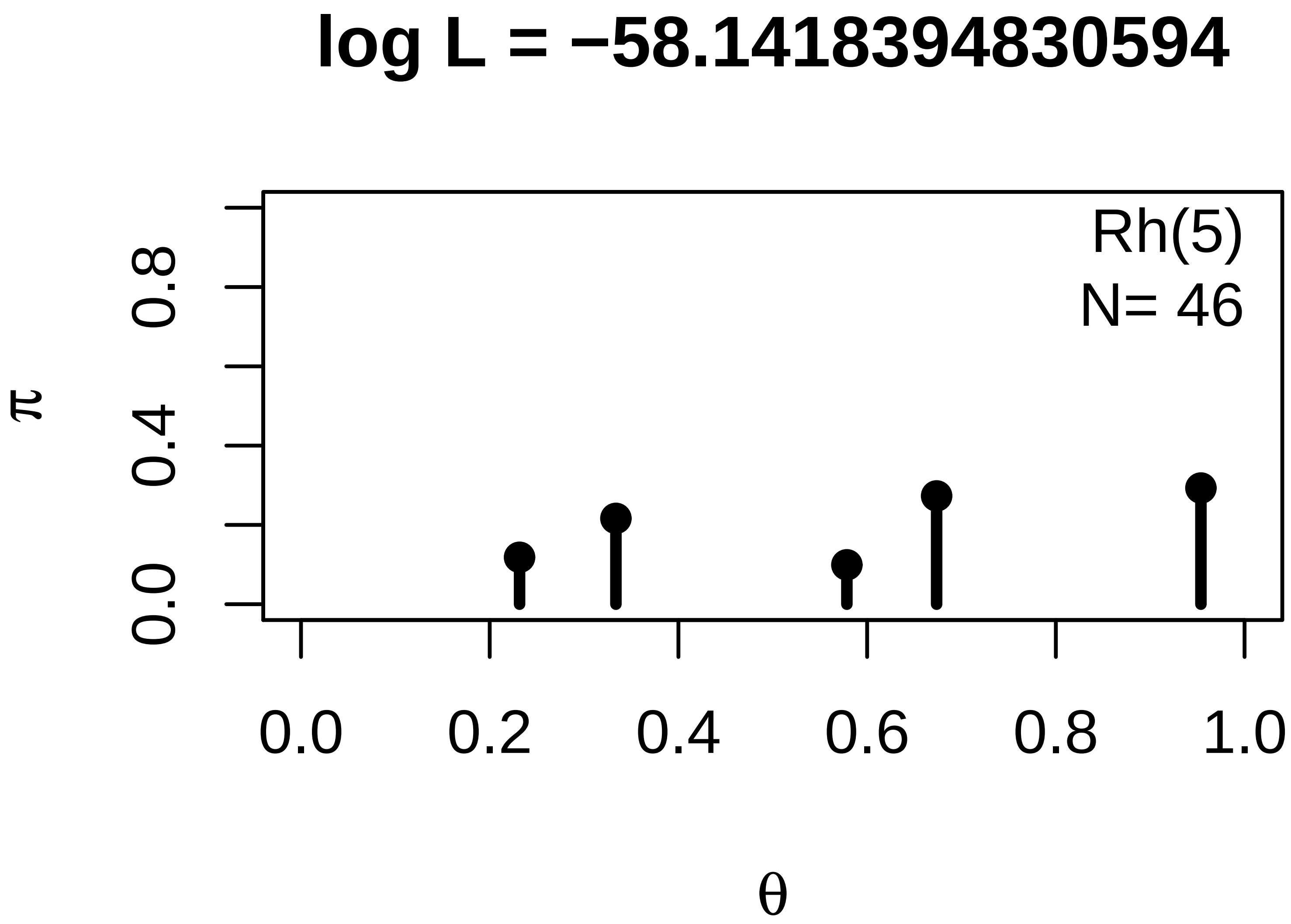
Step 1



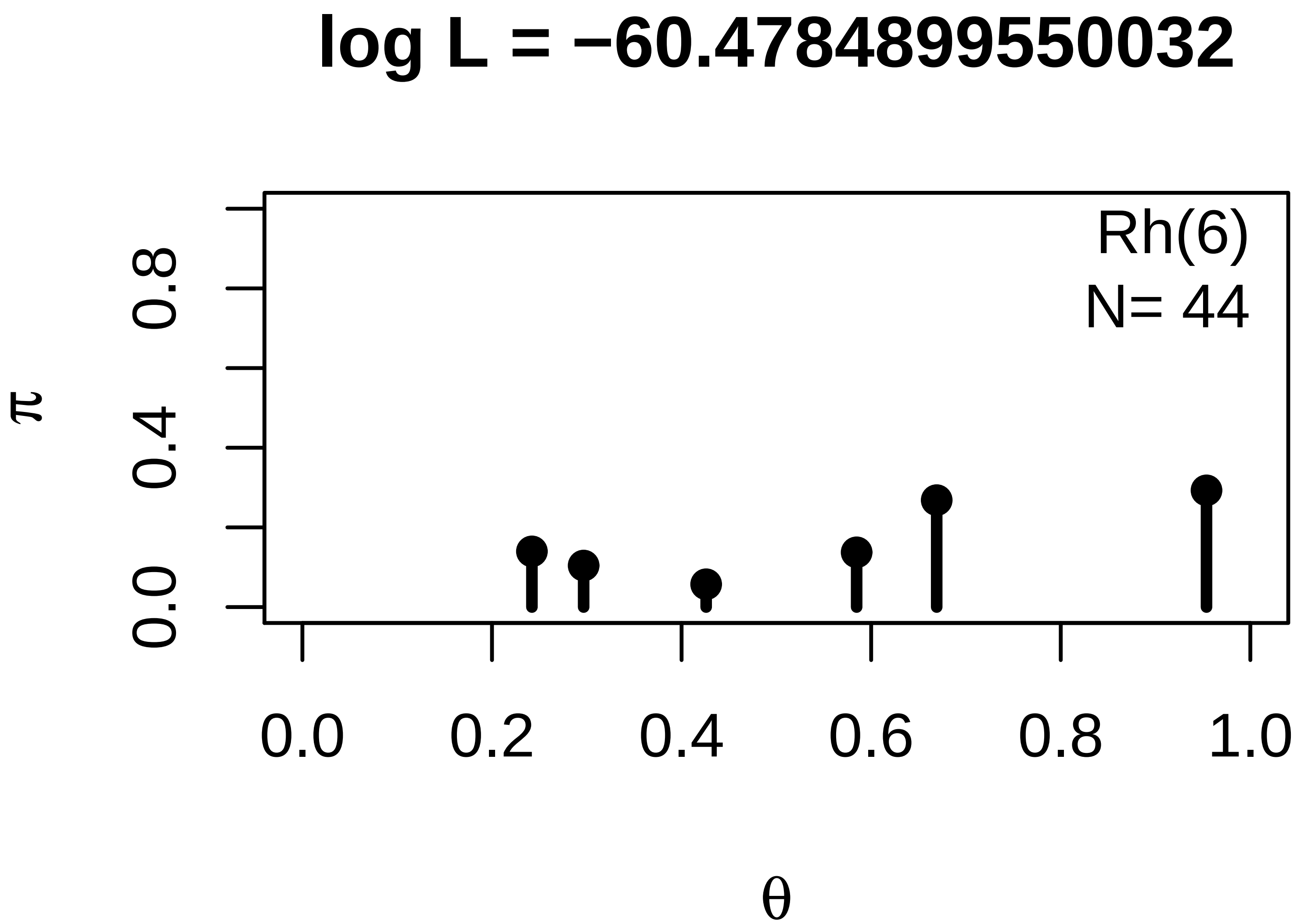
Step 2



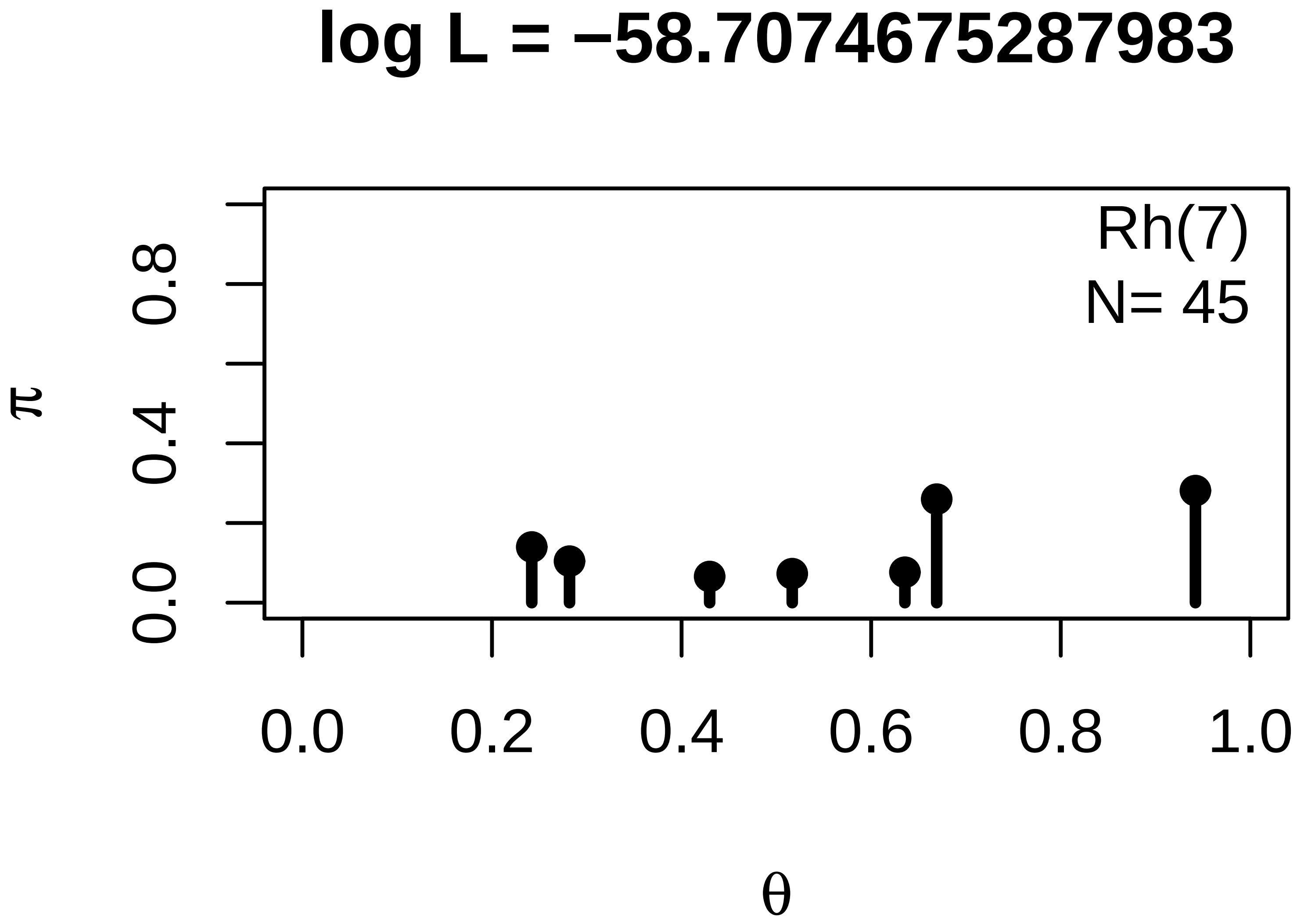
Step 3



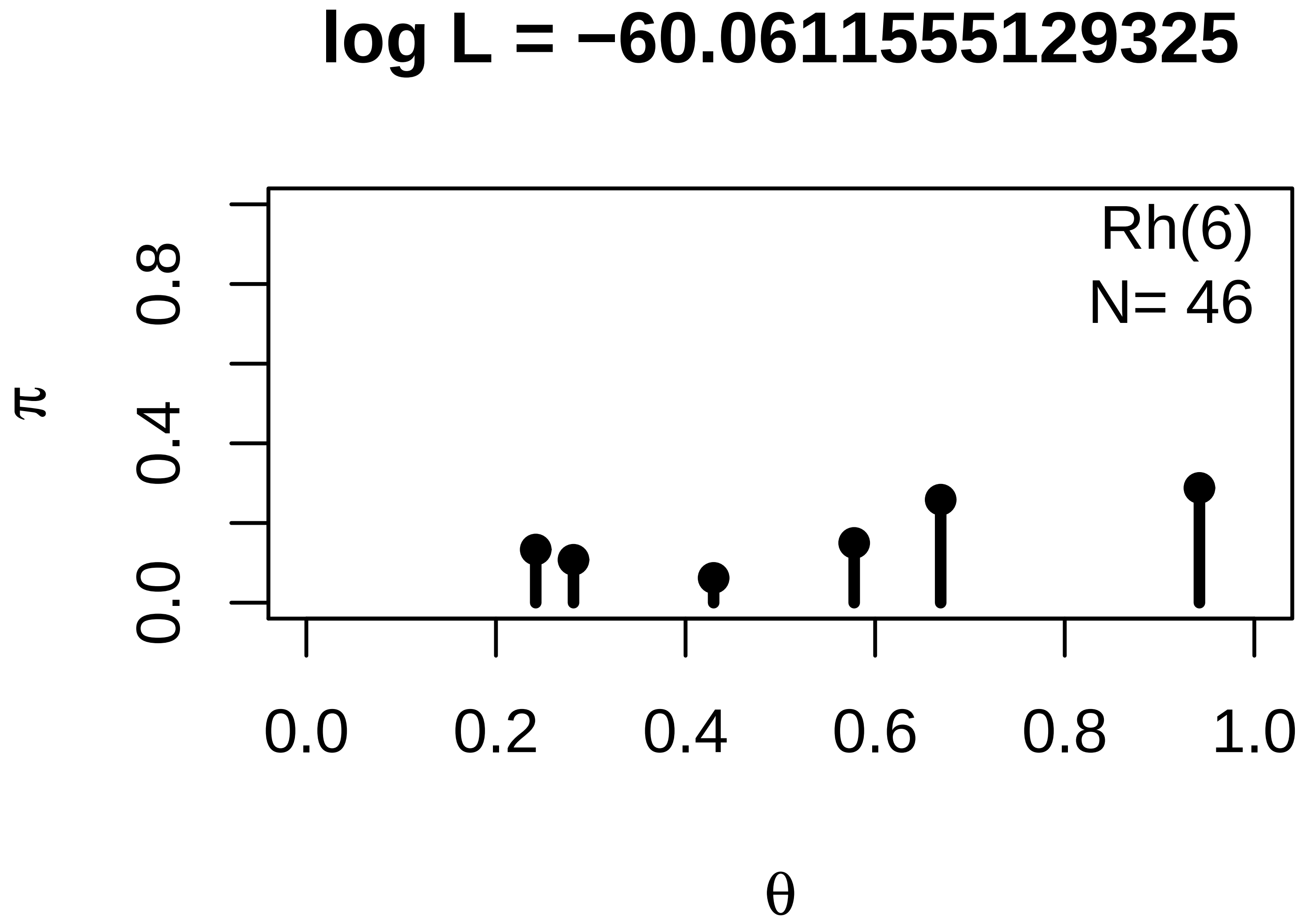
Step 4



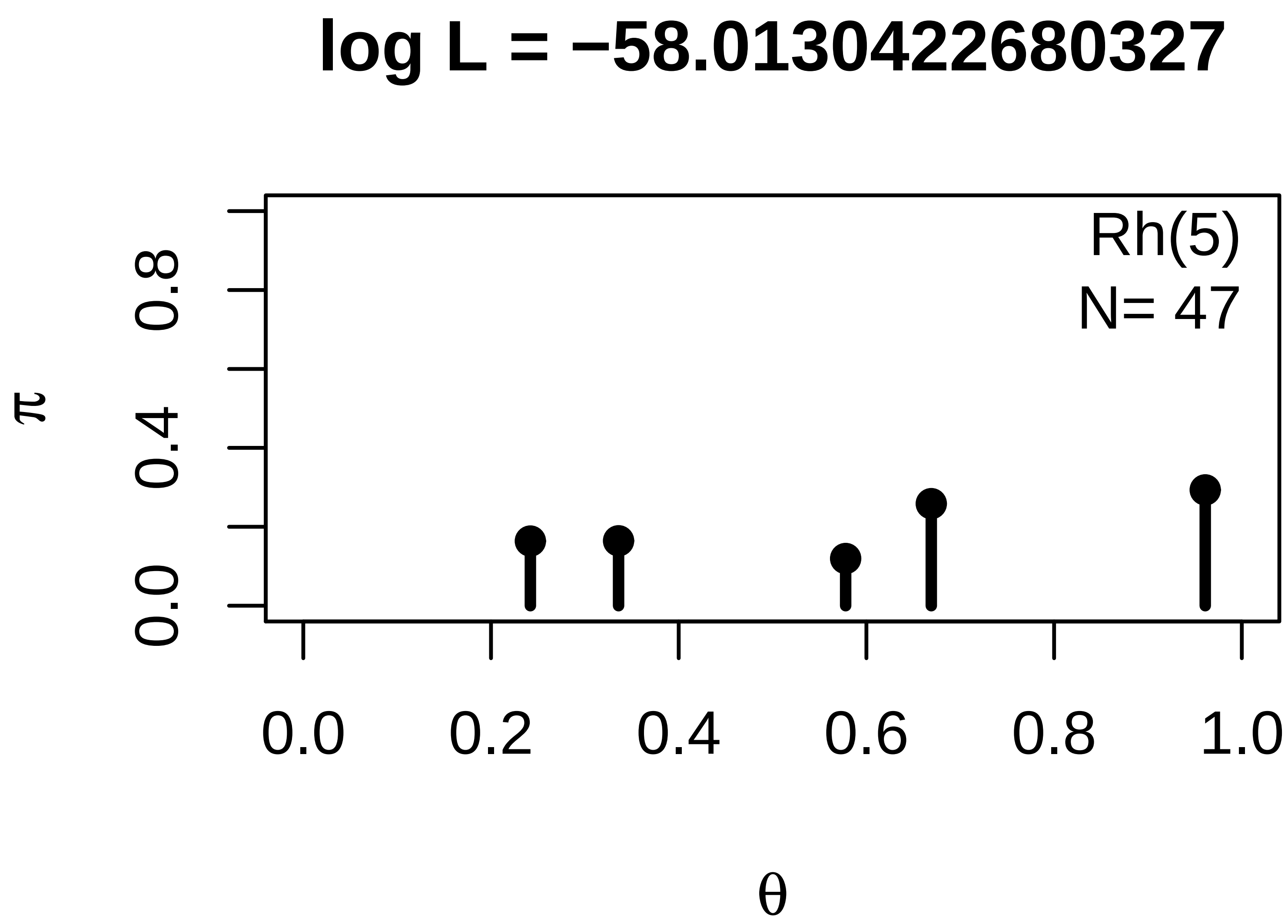
Step 5



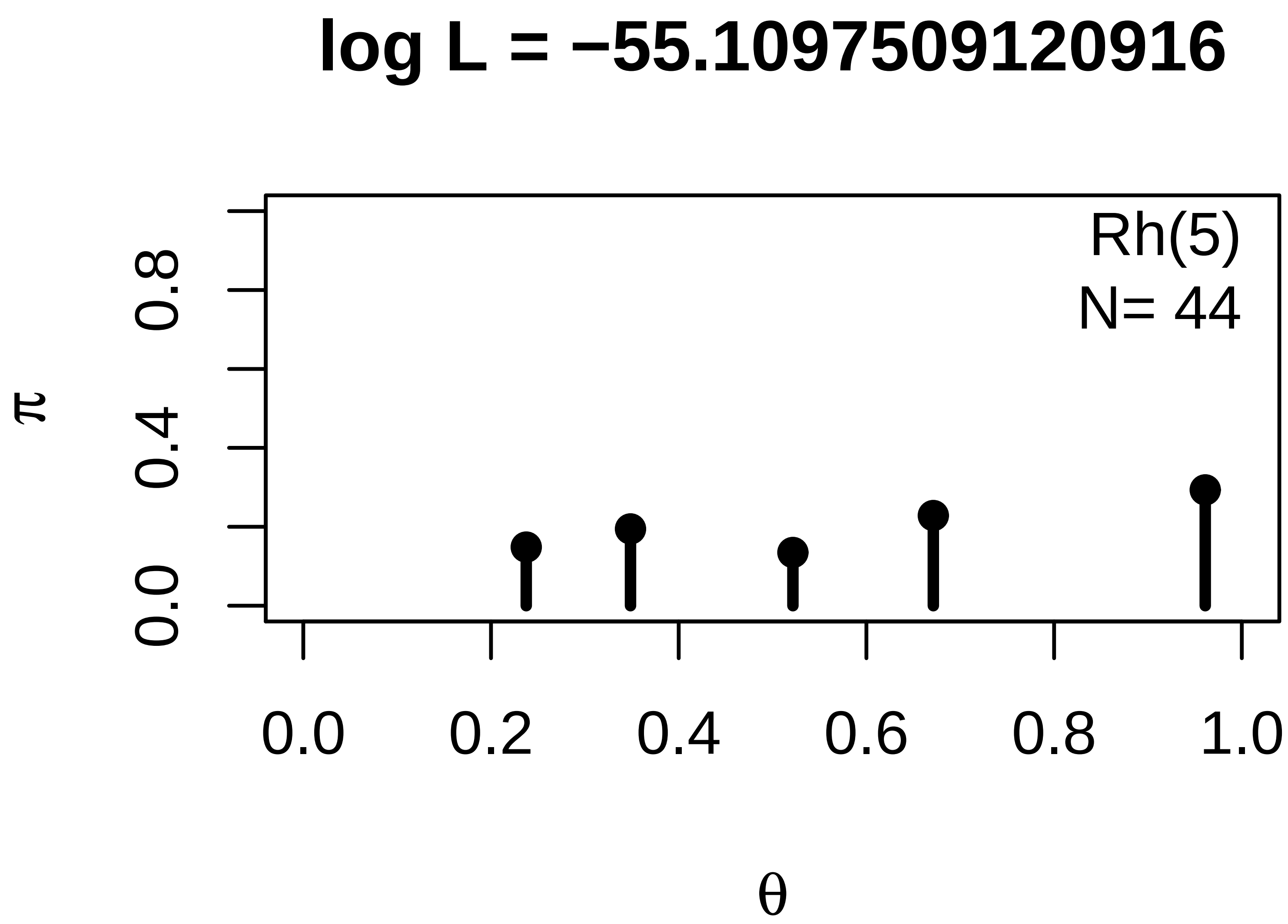
Step 6



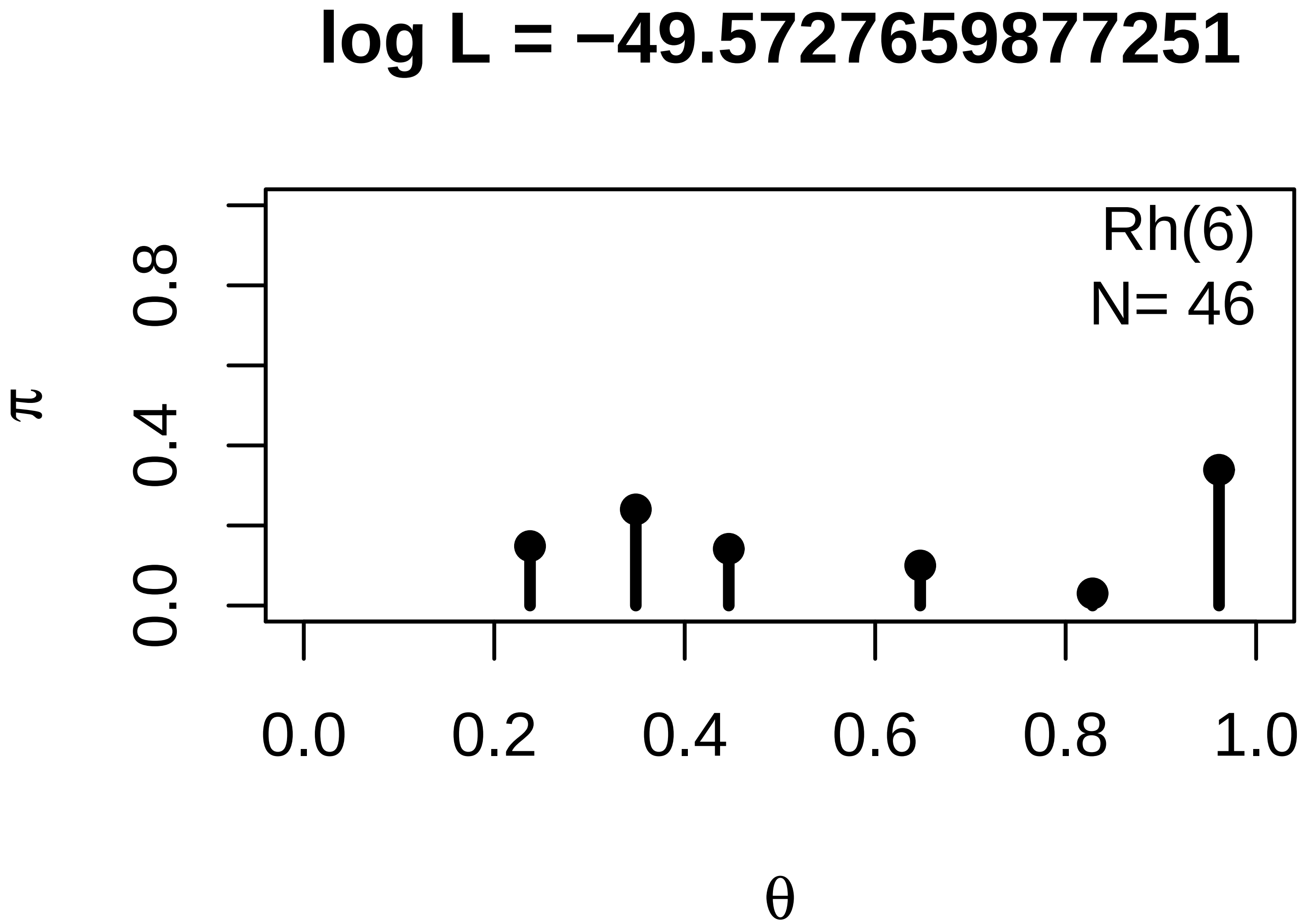
Step 7



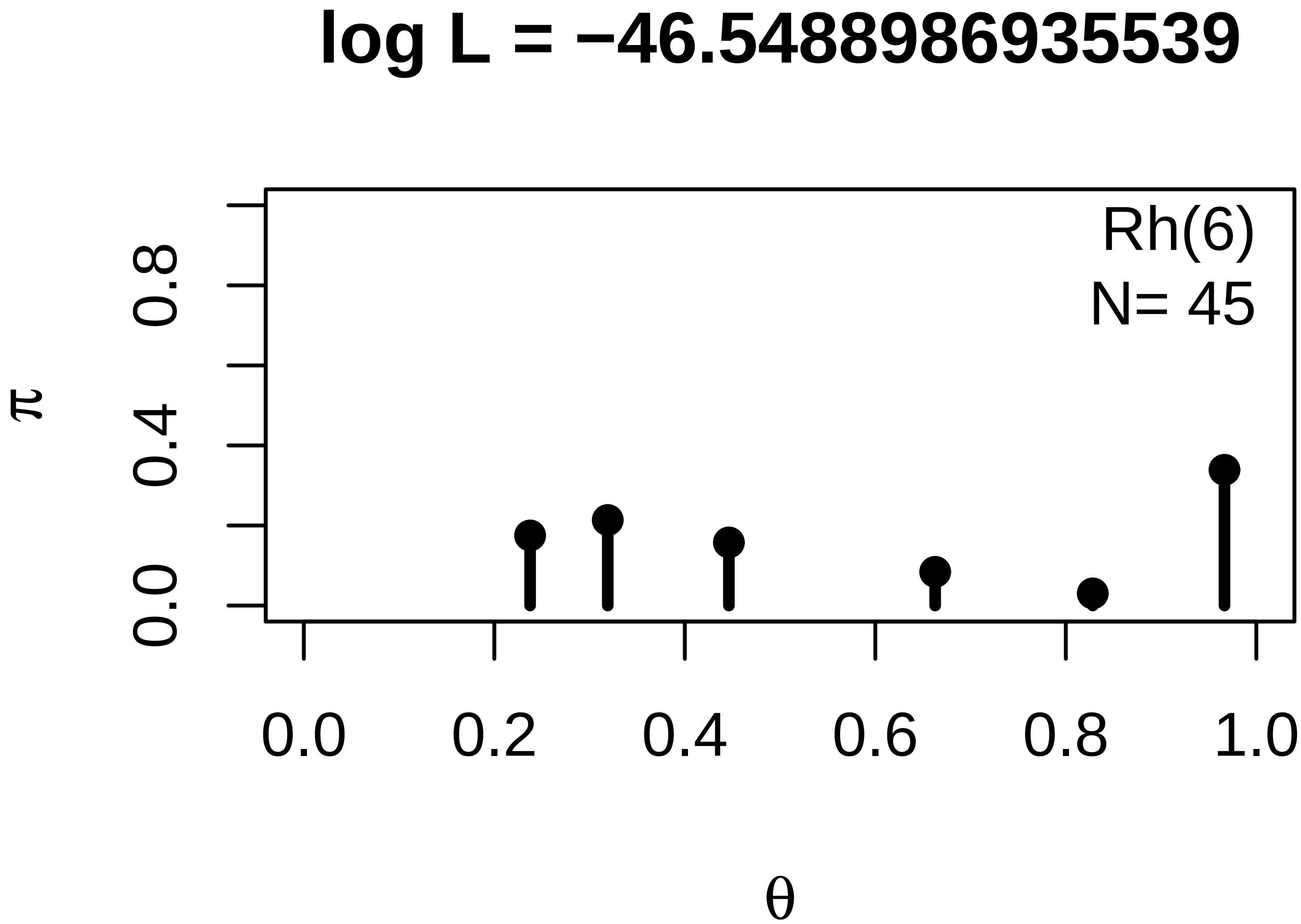
Step 8



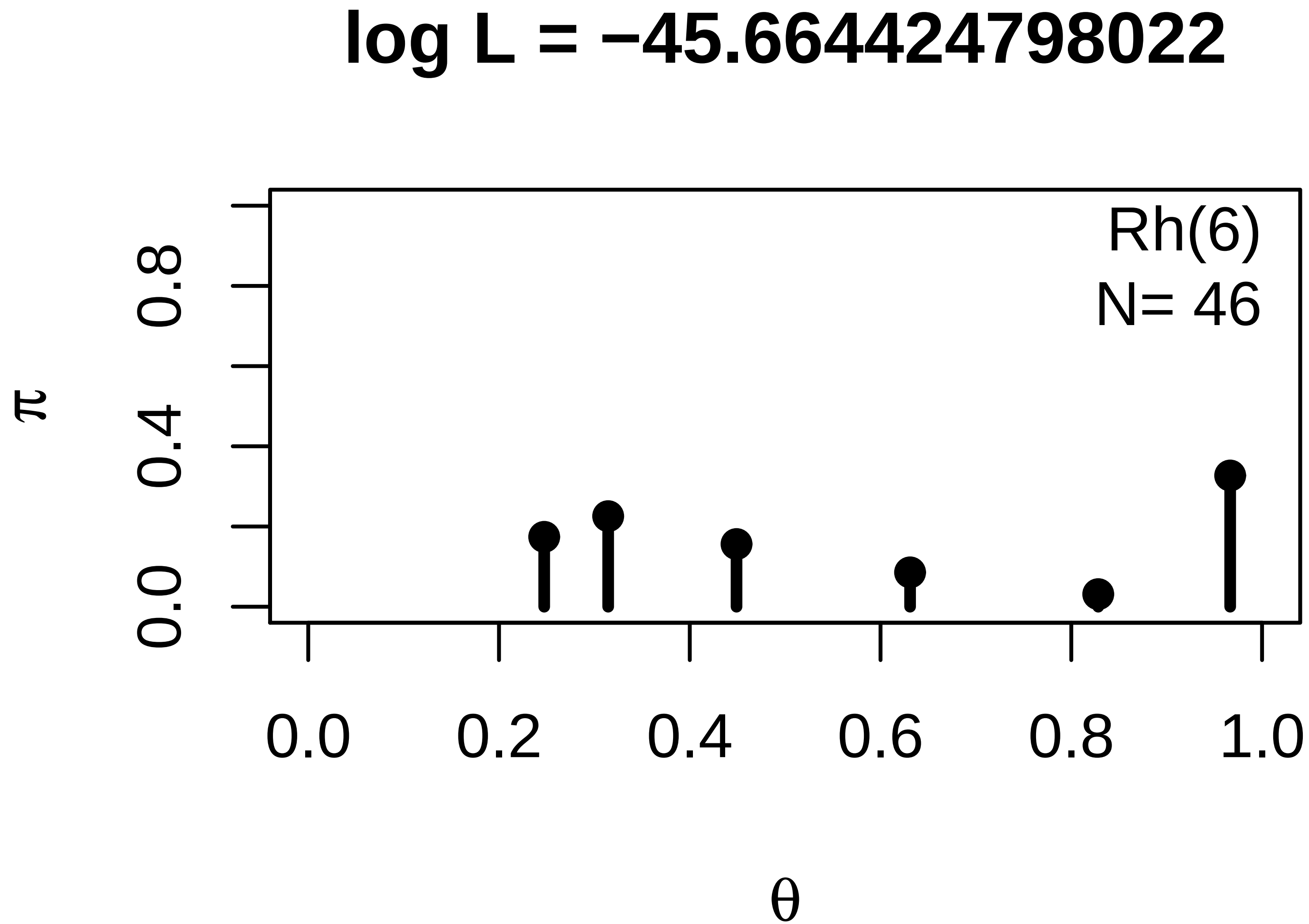
Step 9



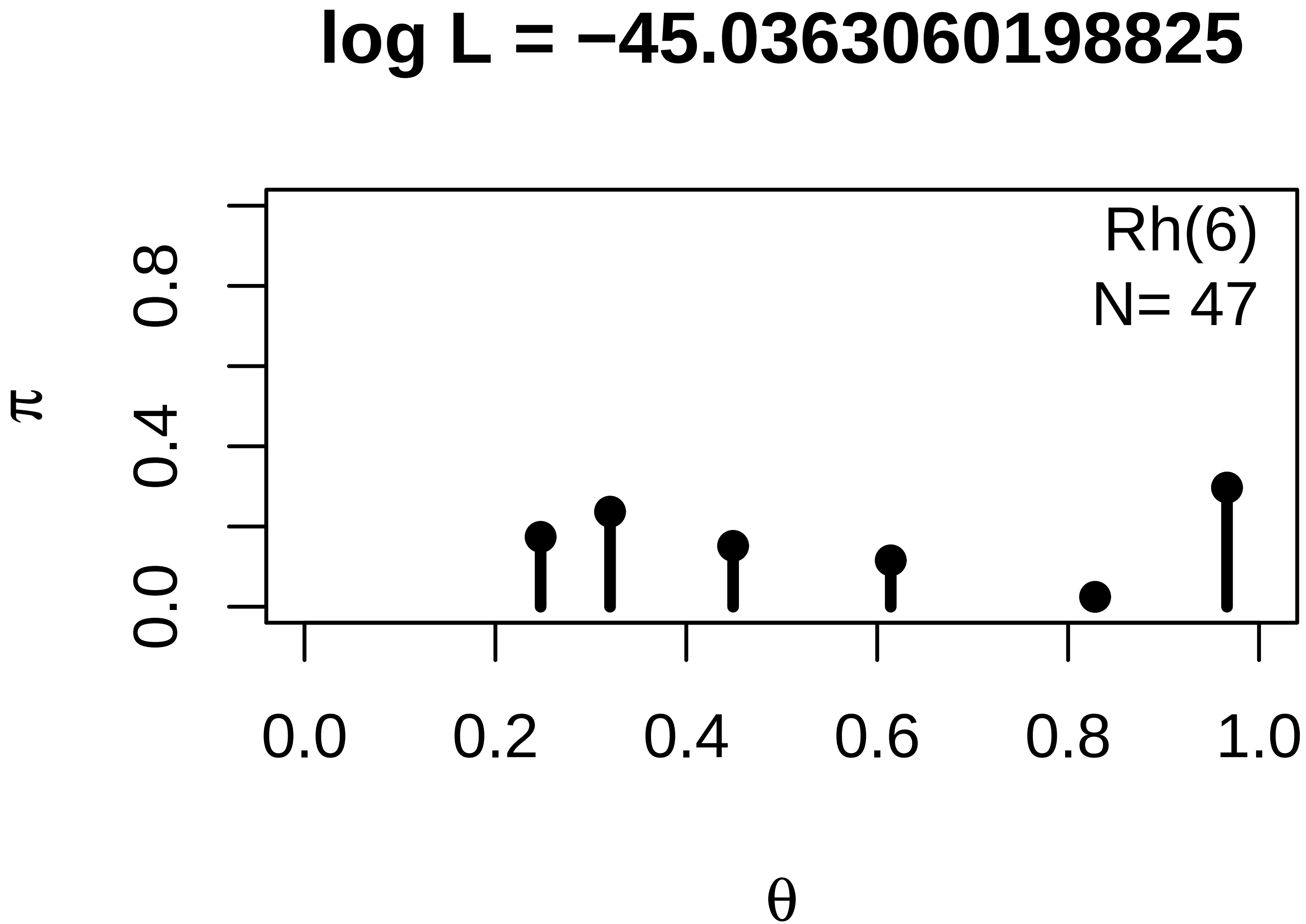
Step 10



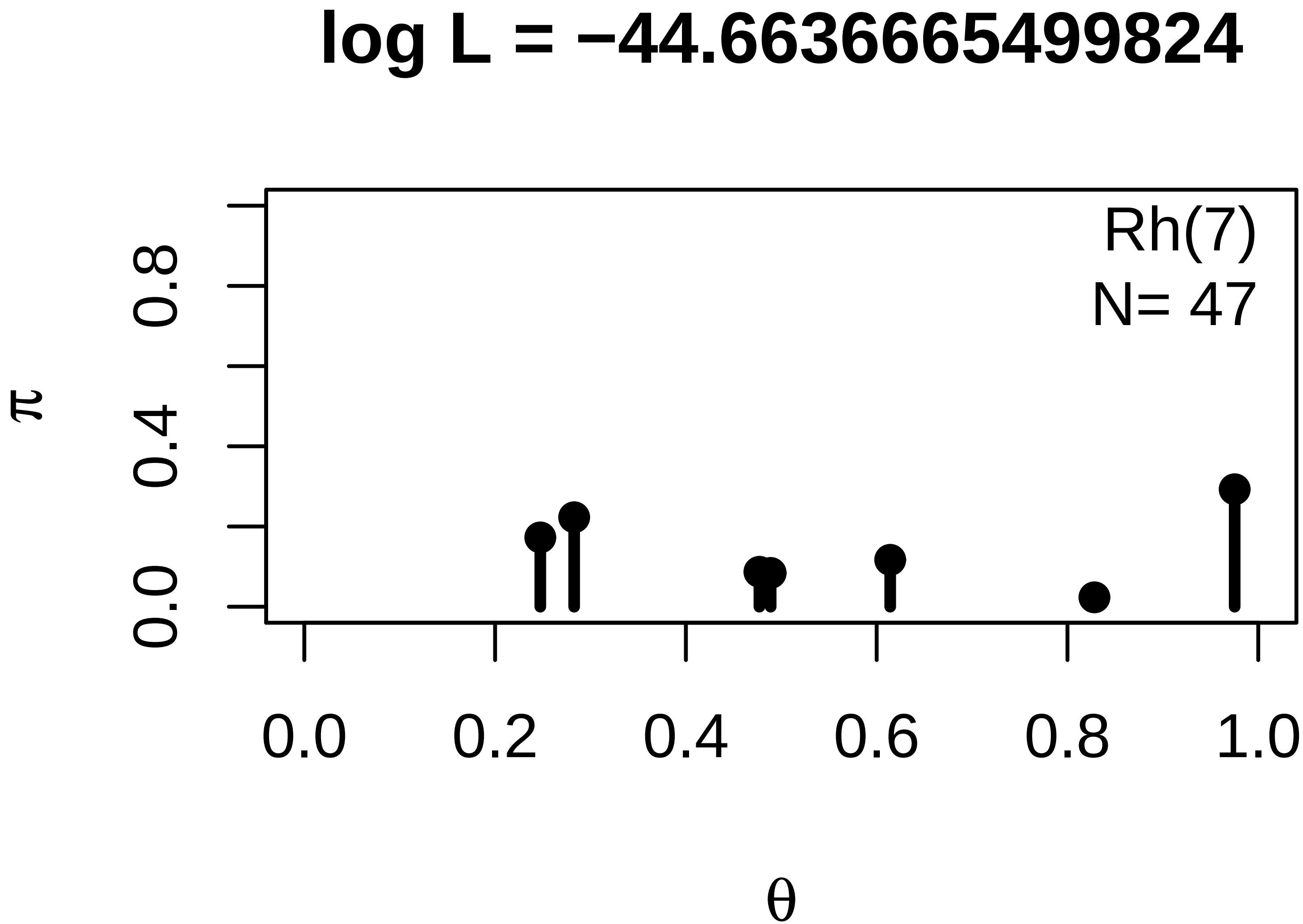
Step 11



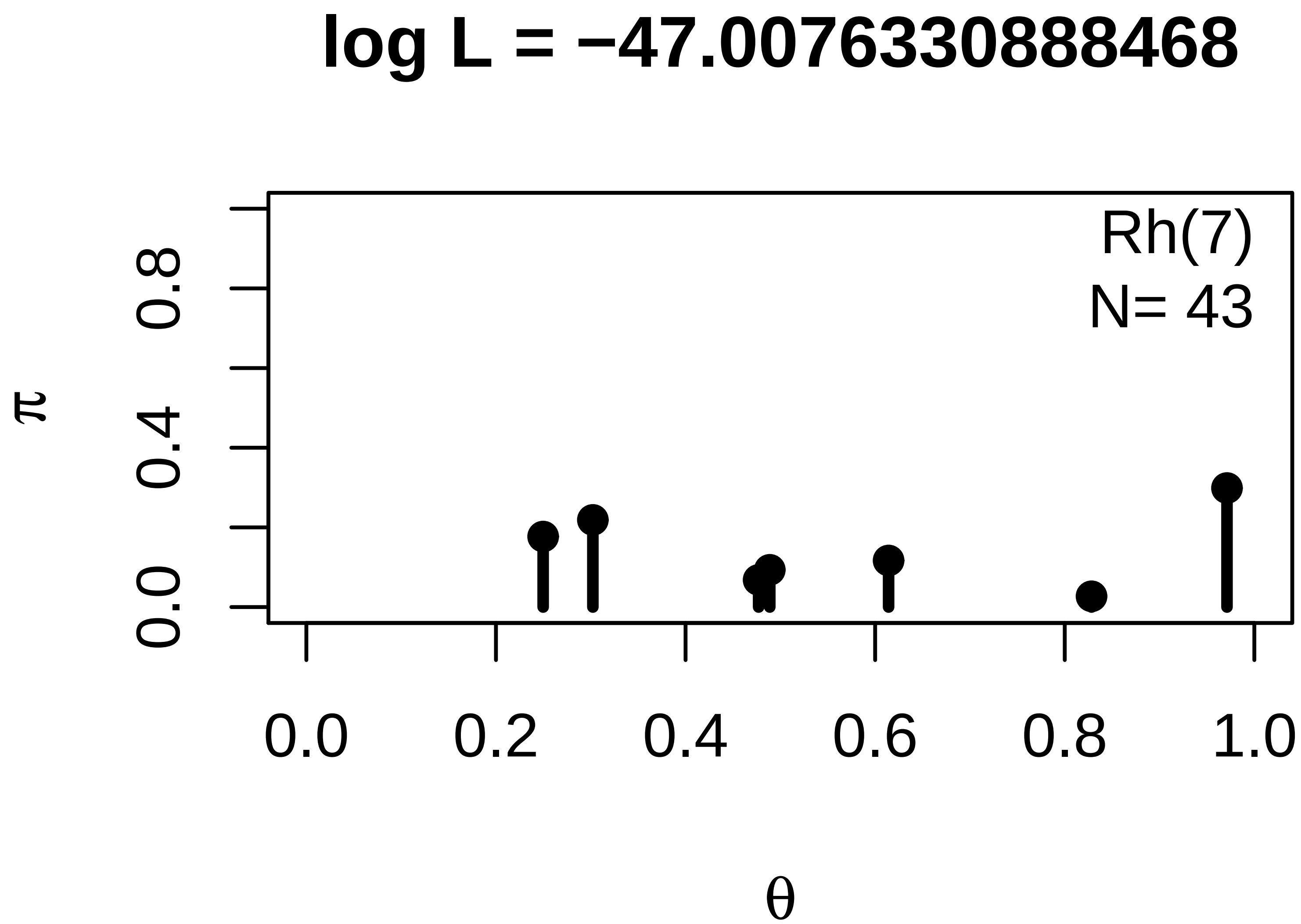
Step 12



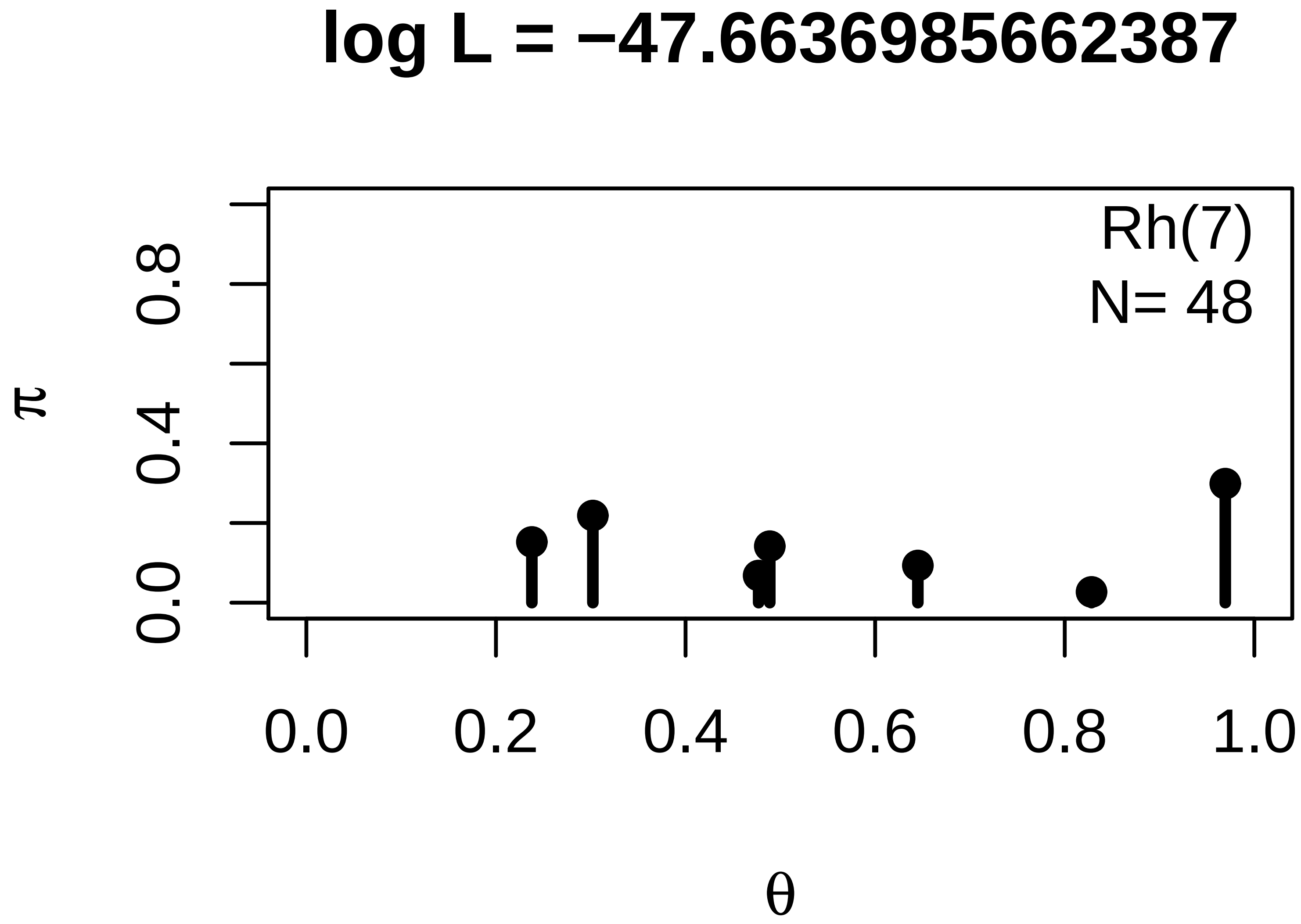
Step 13



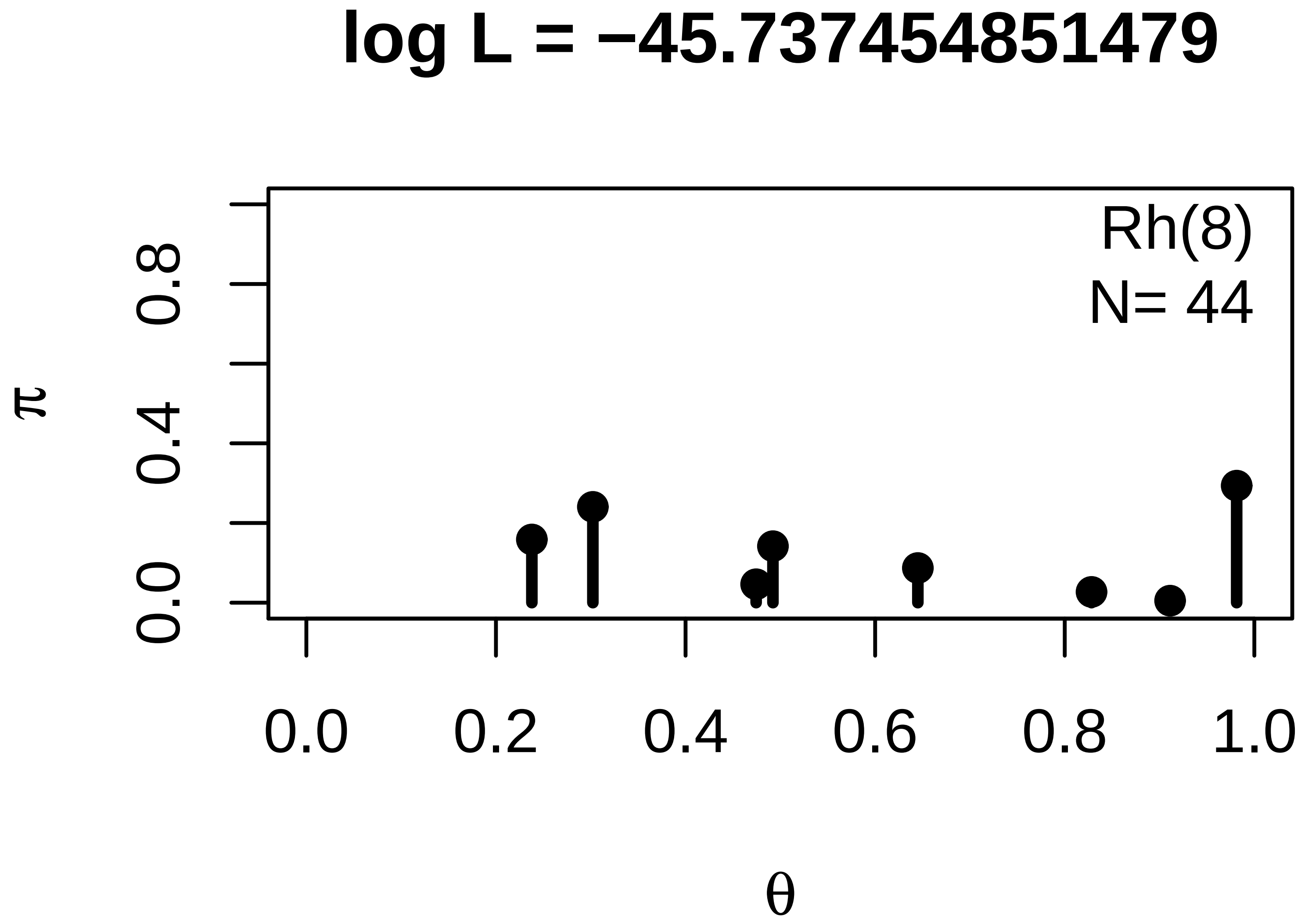
Step 14



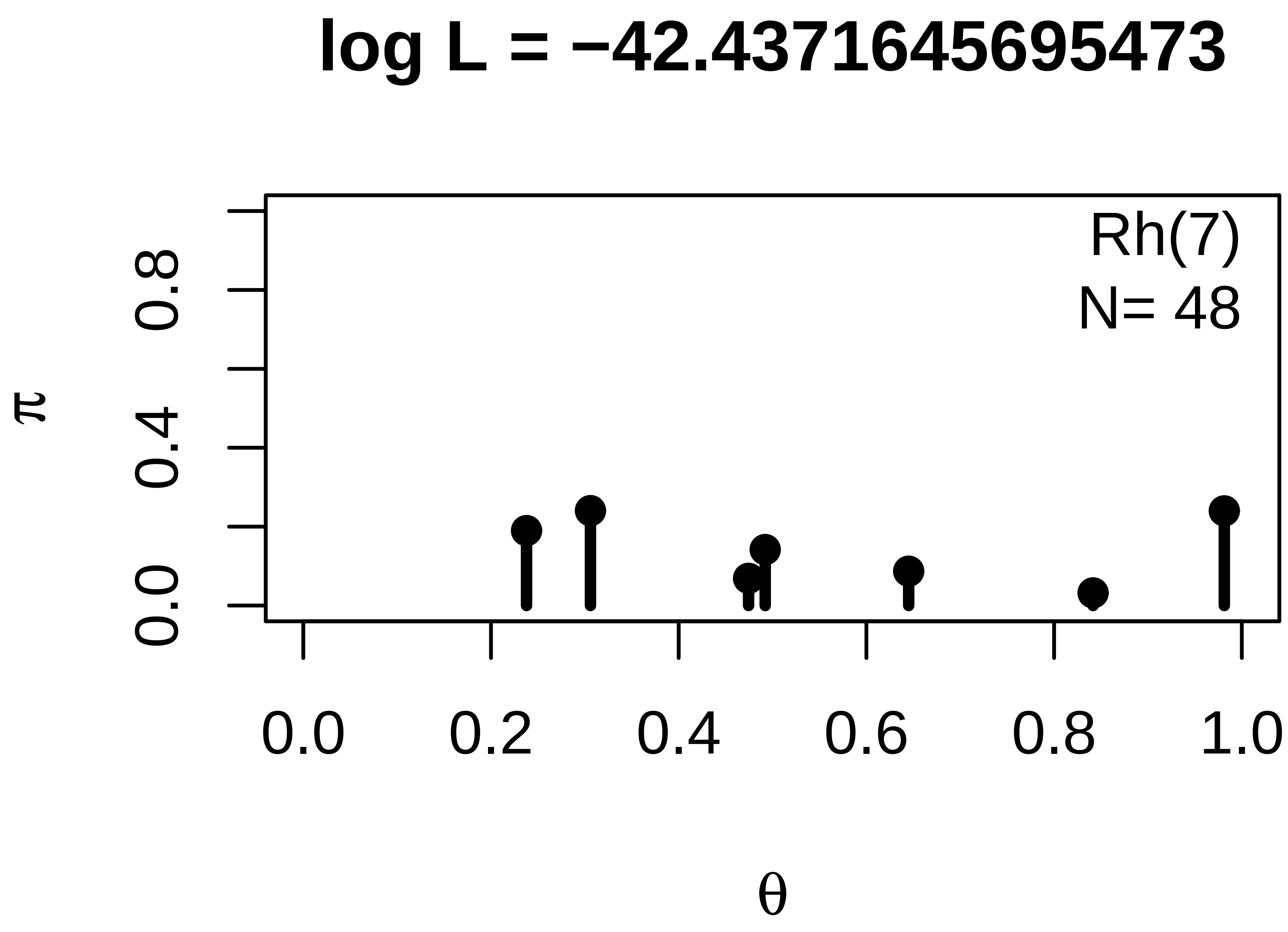
Step 15



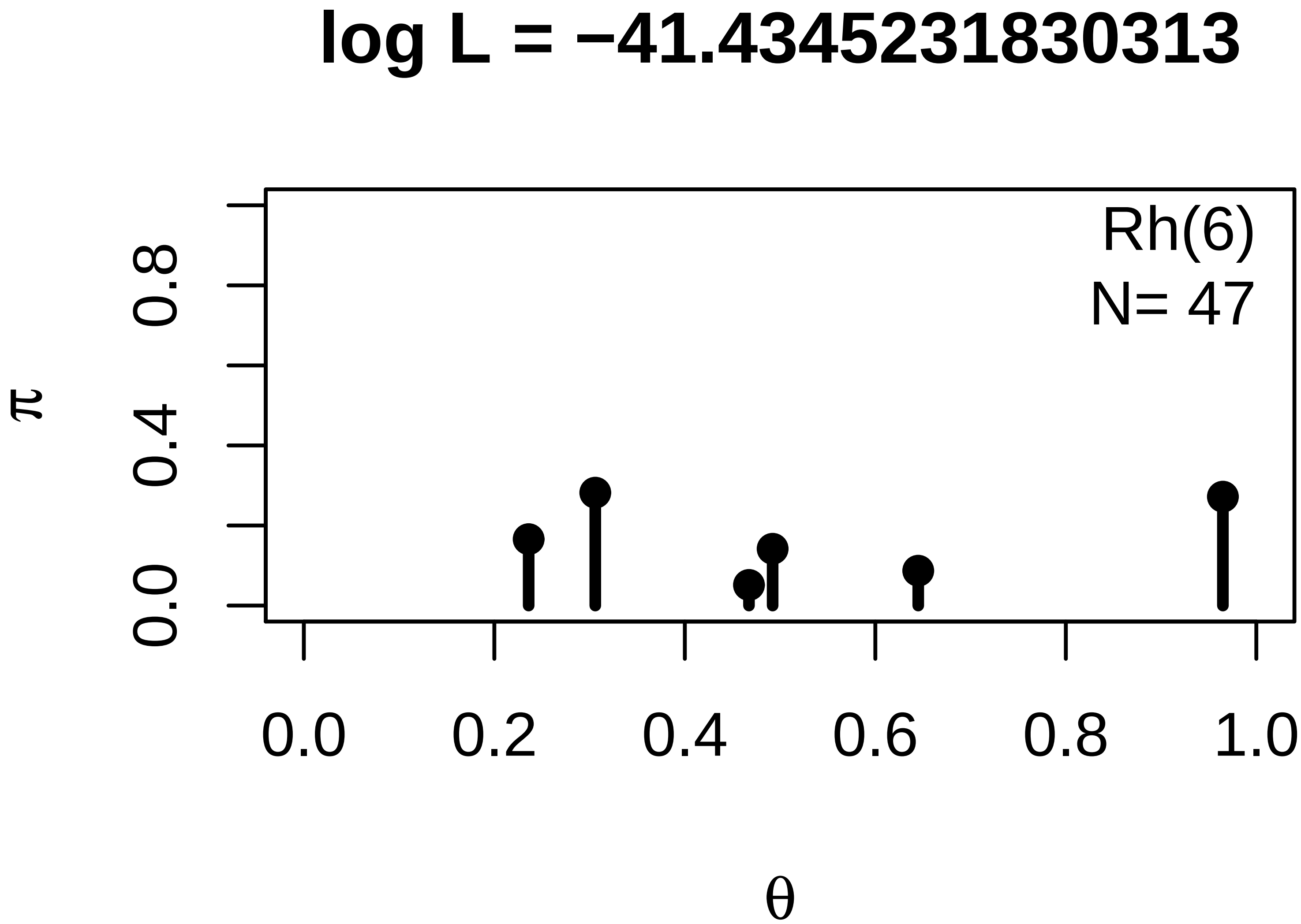
Step 16



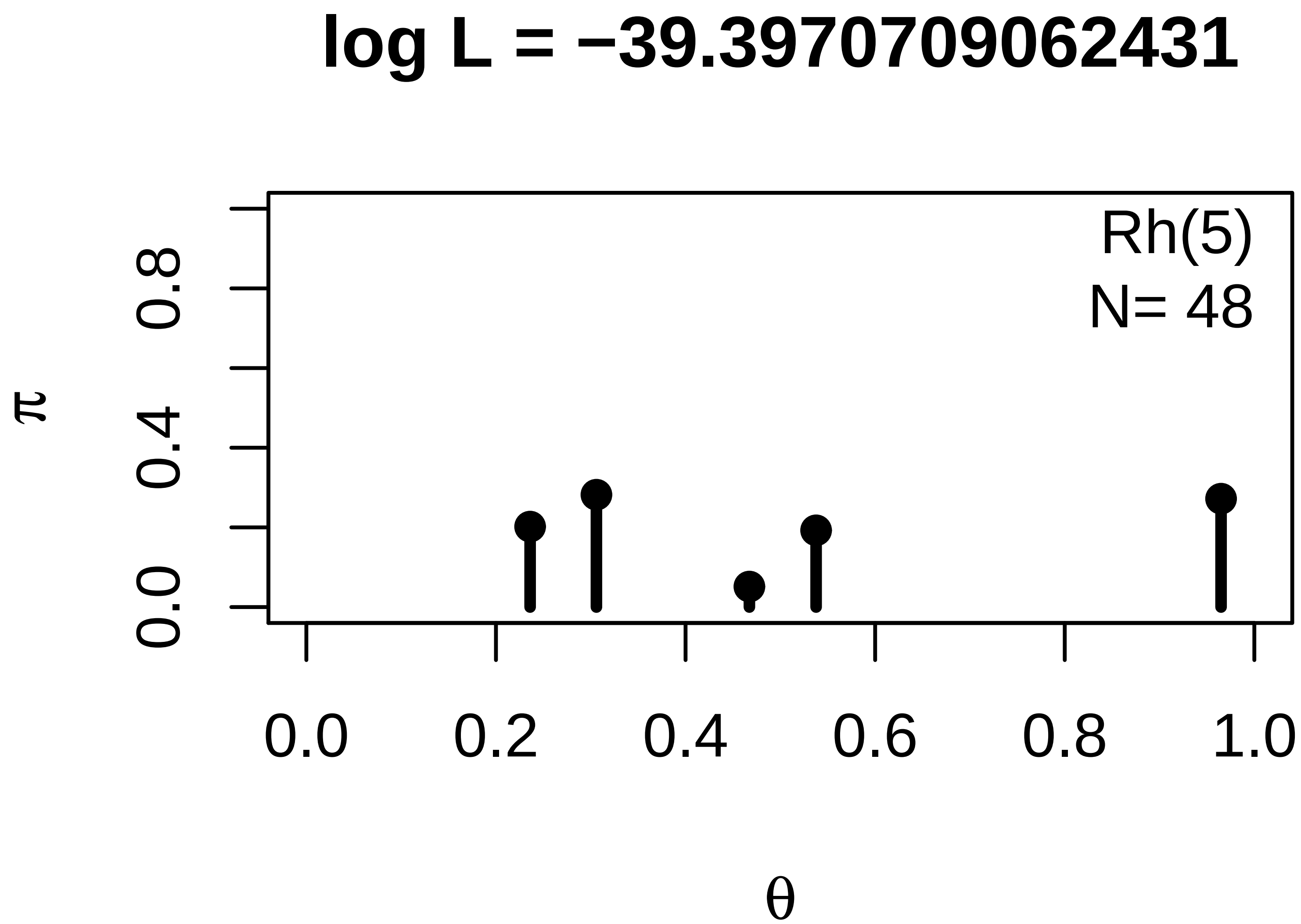
Step 17



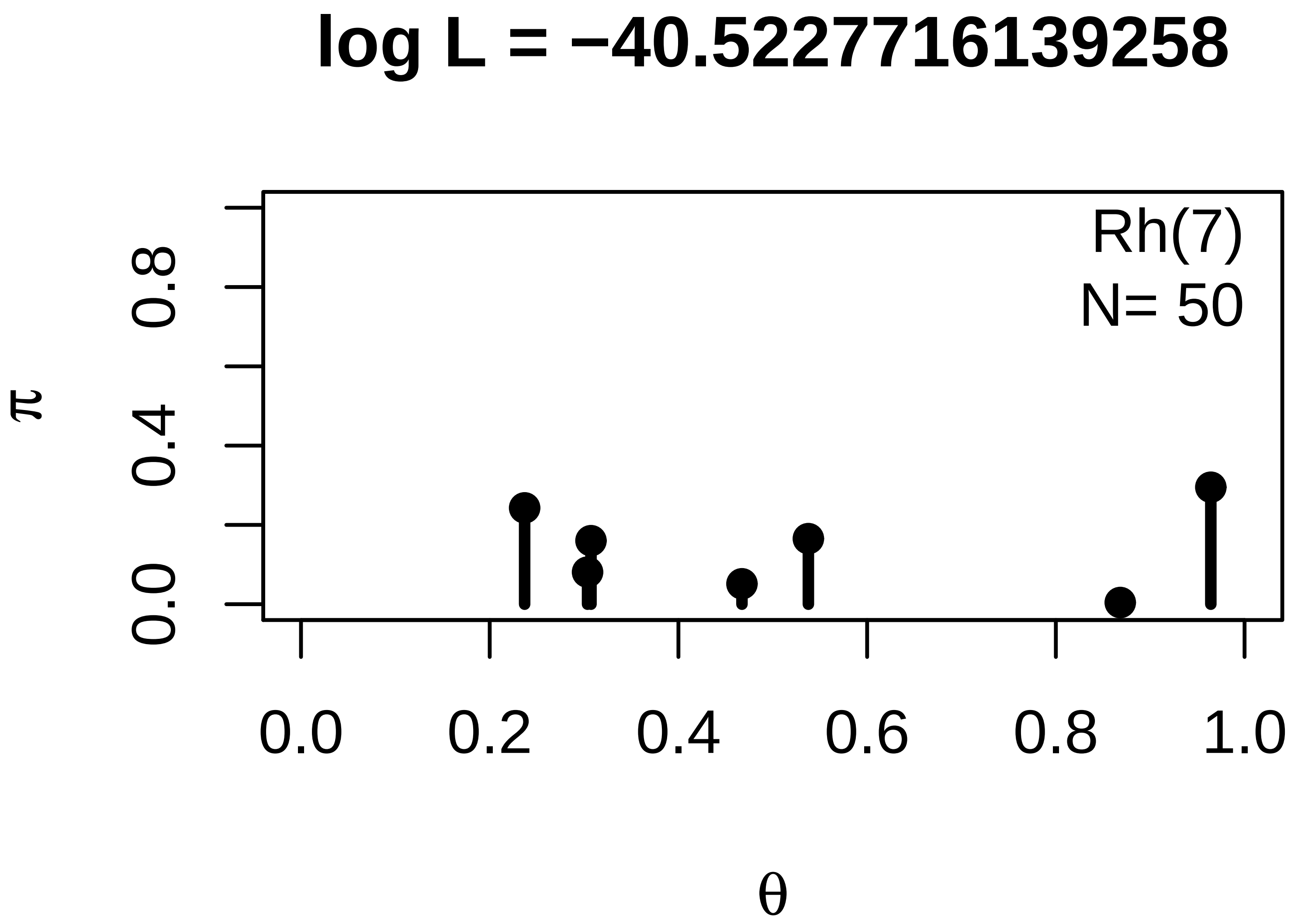
Step 18



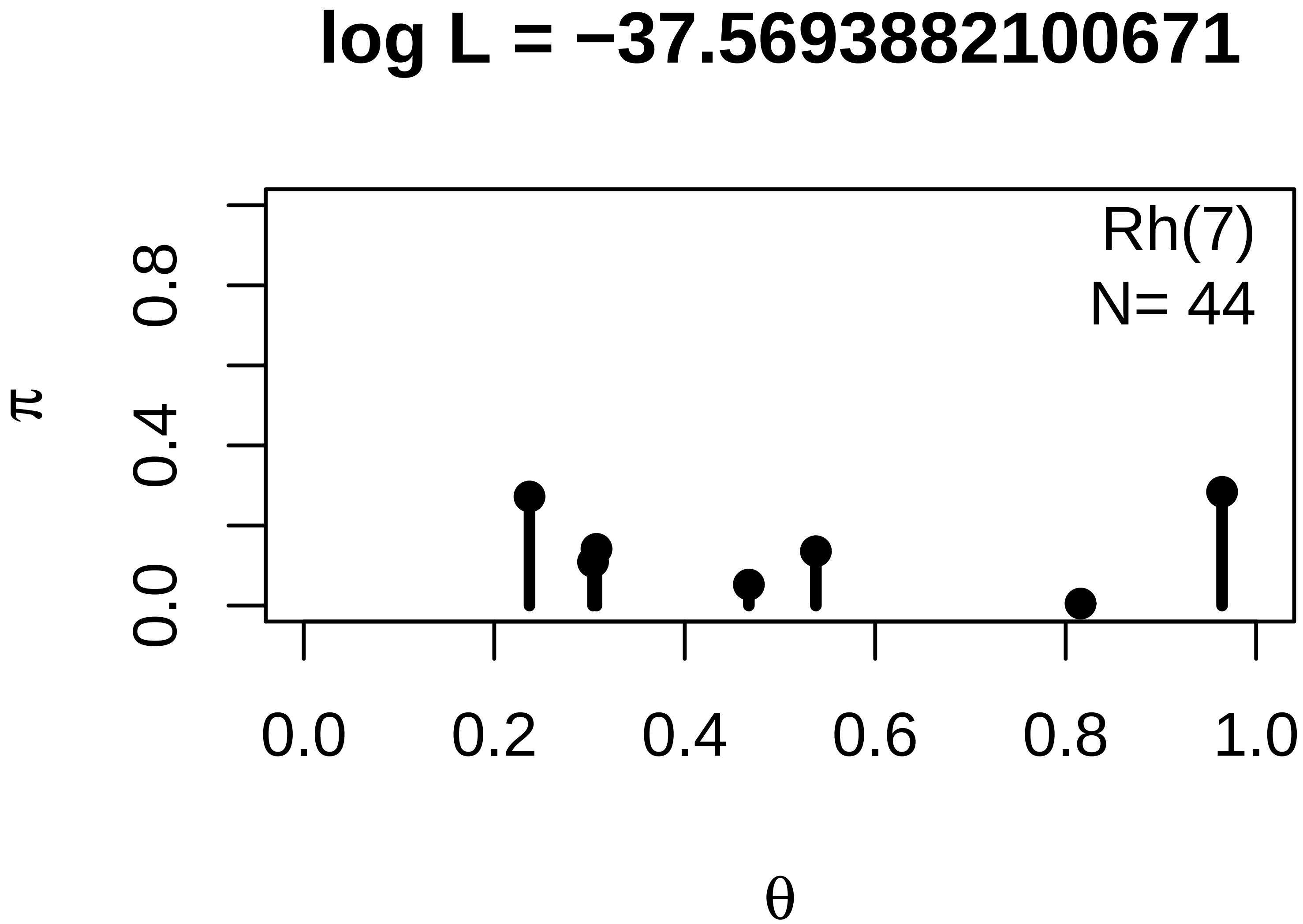
Step 19



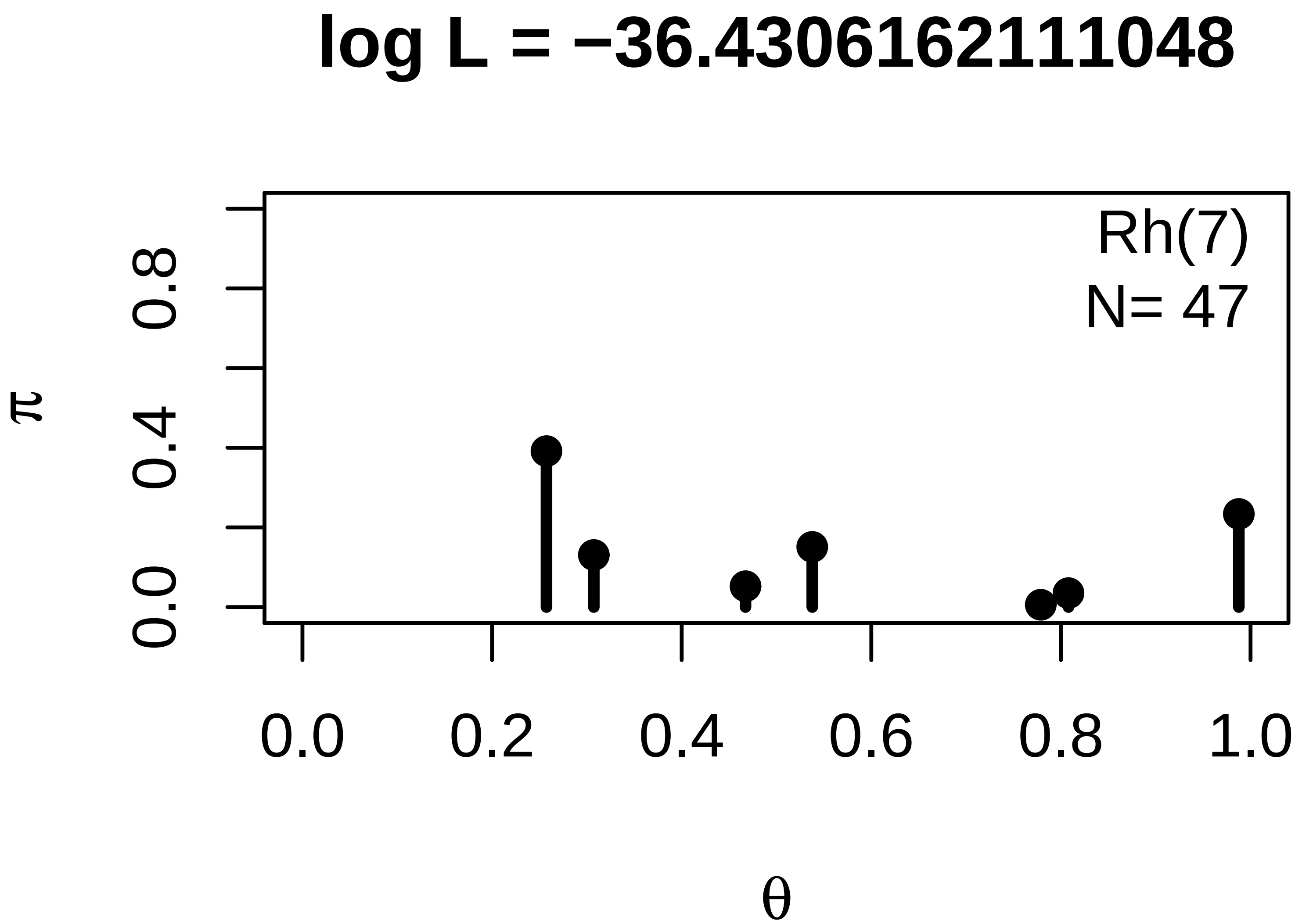
Step 20



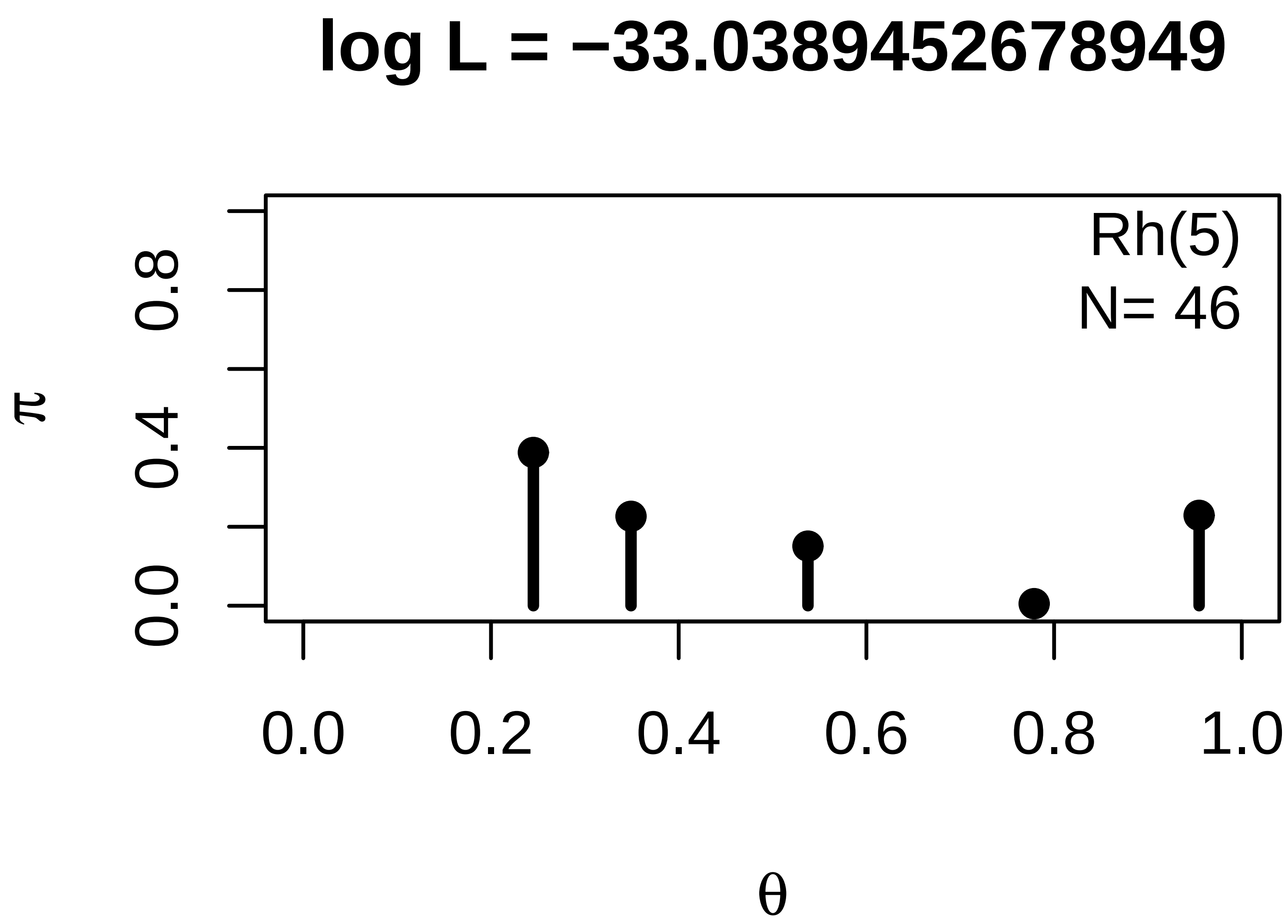
Step 21



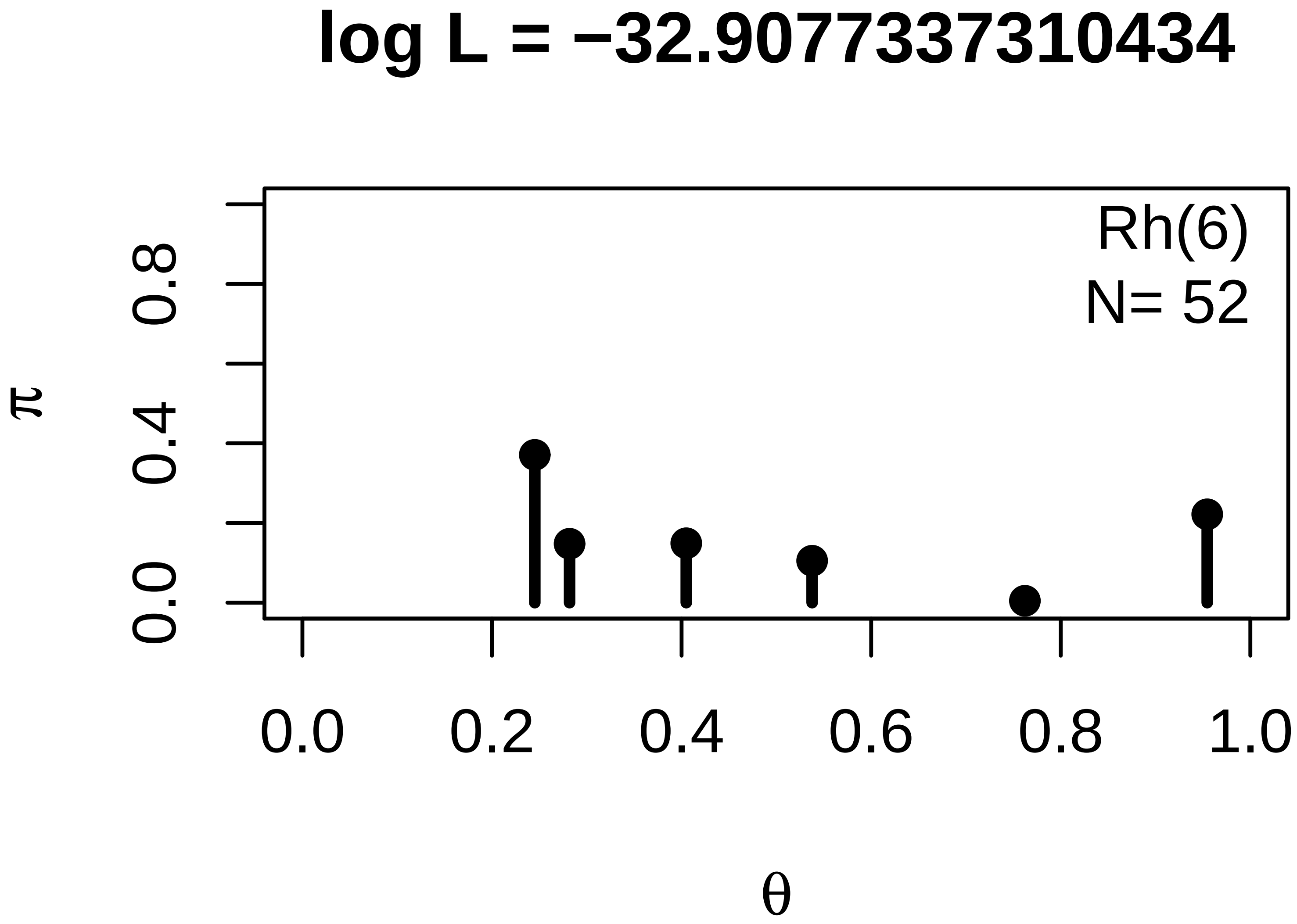
Step 22



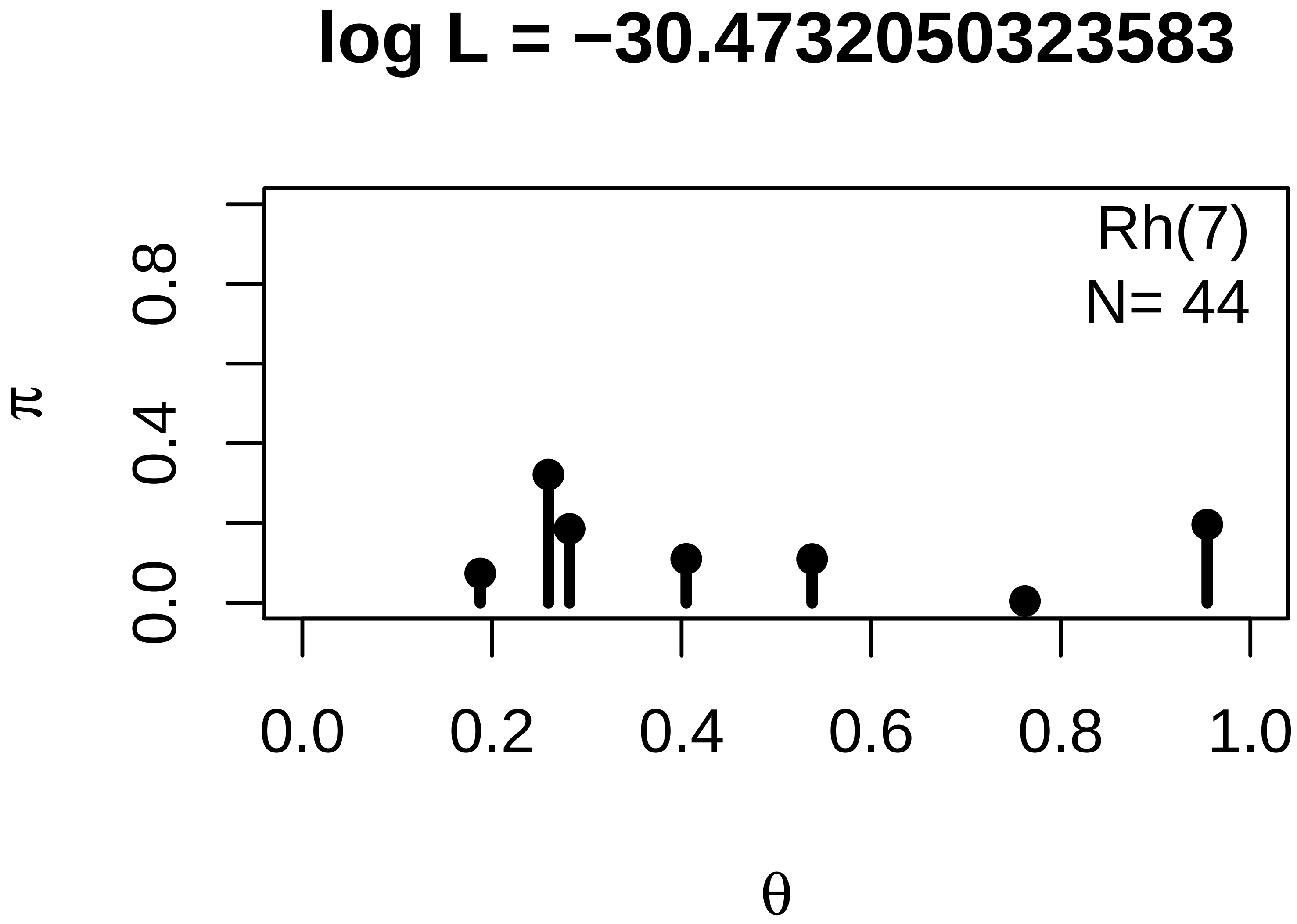
Step 23



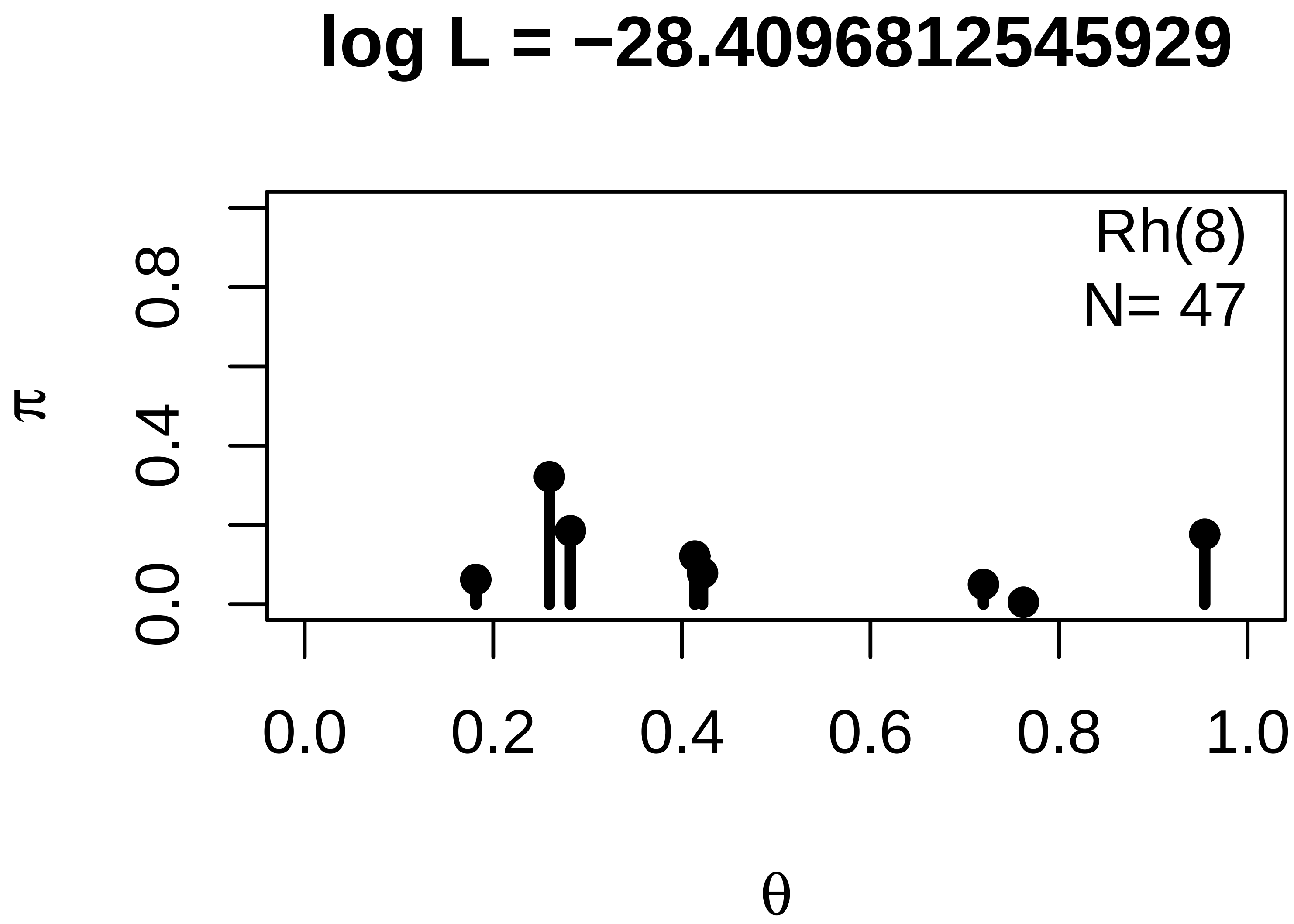
Step 24



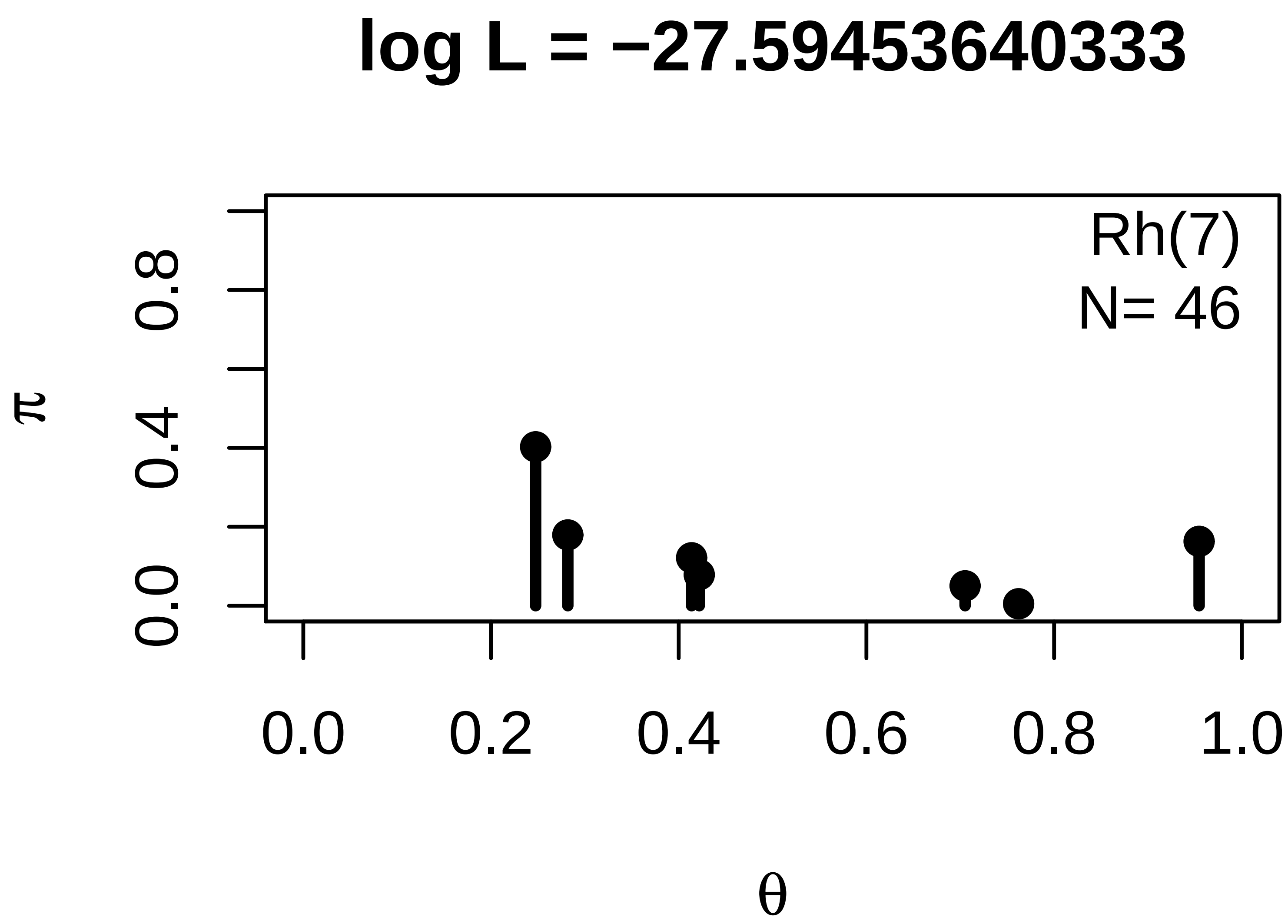
Step 25



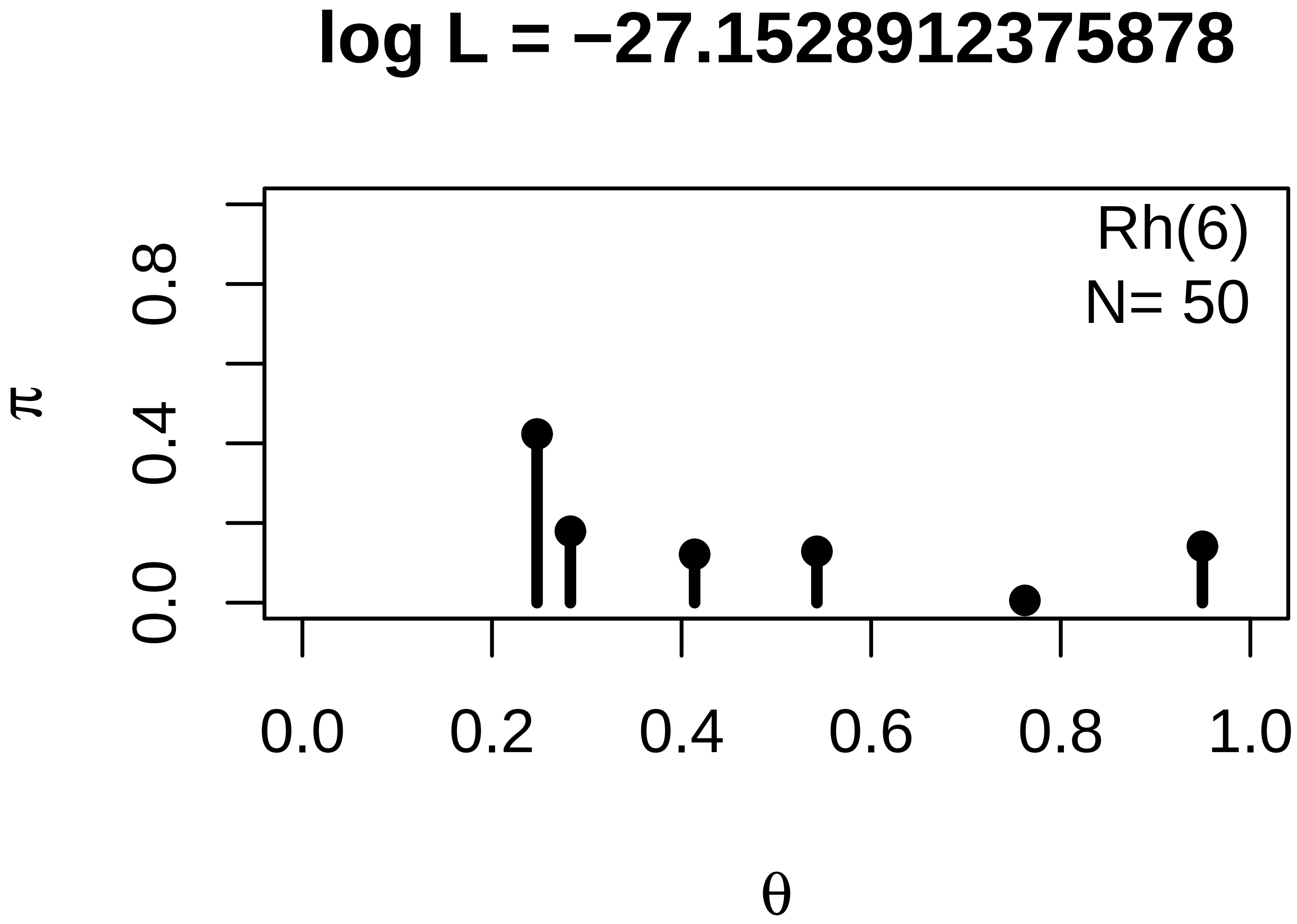
Step 26



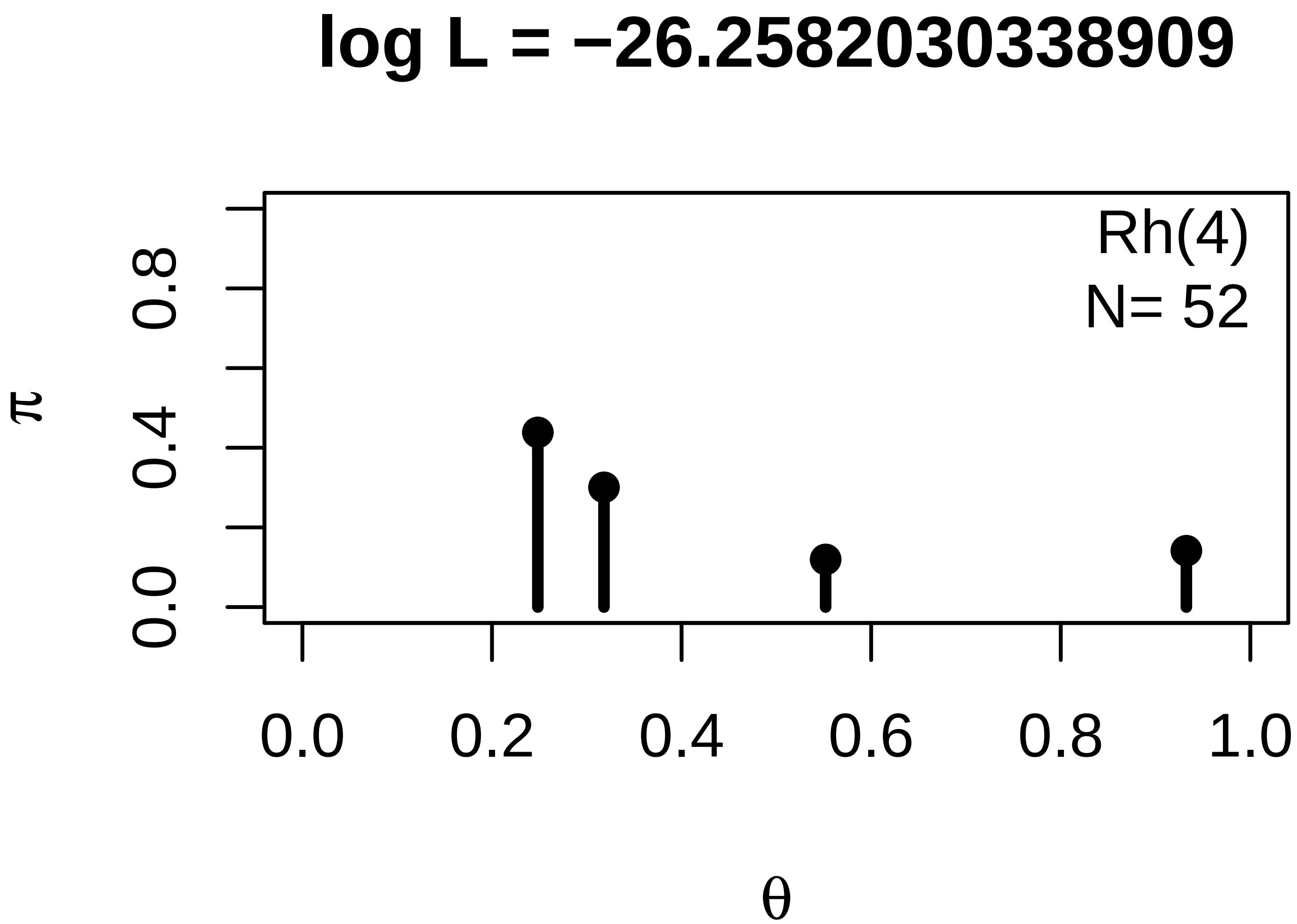
Step 27



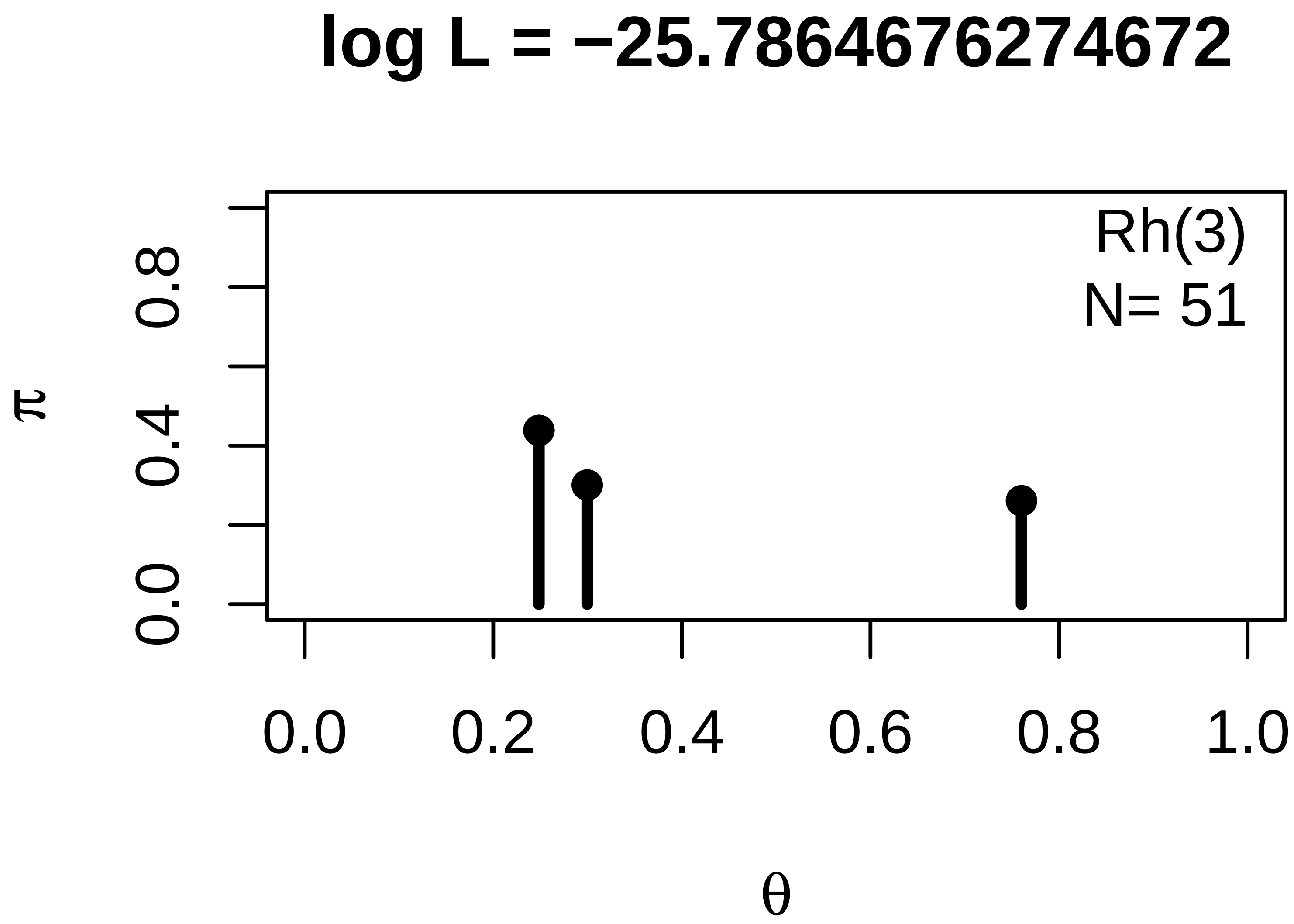
Step 28



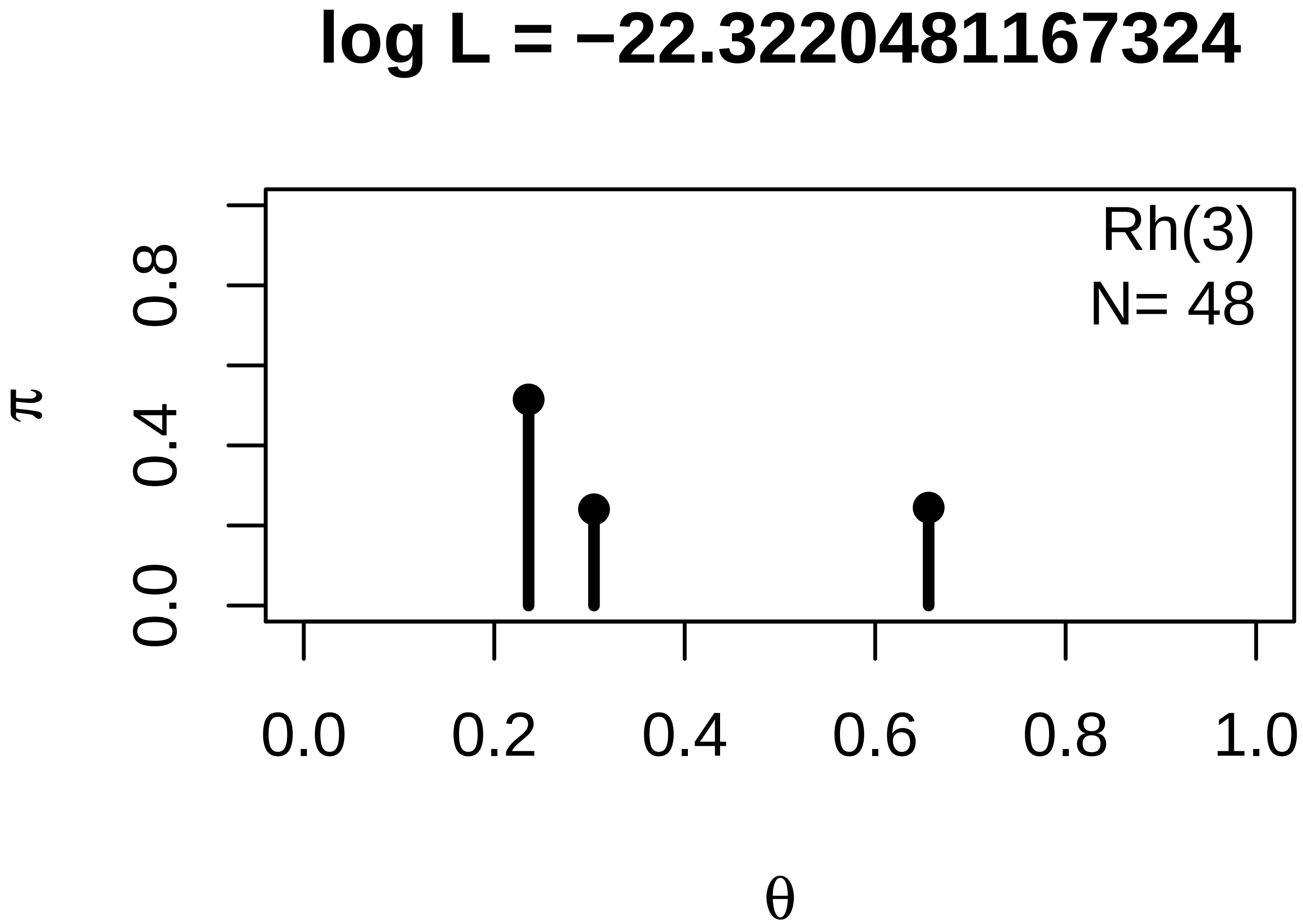
Step 29



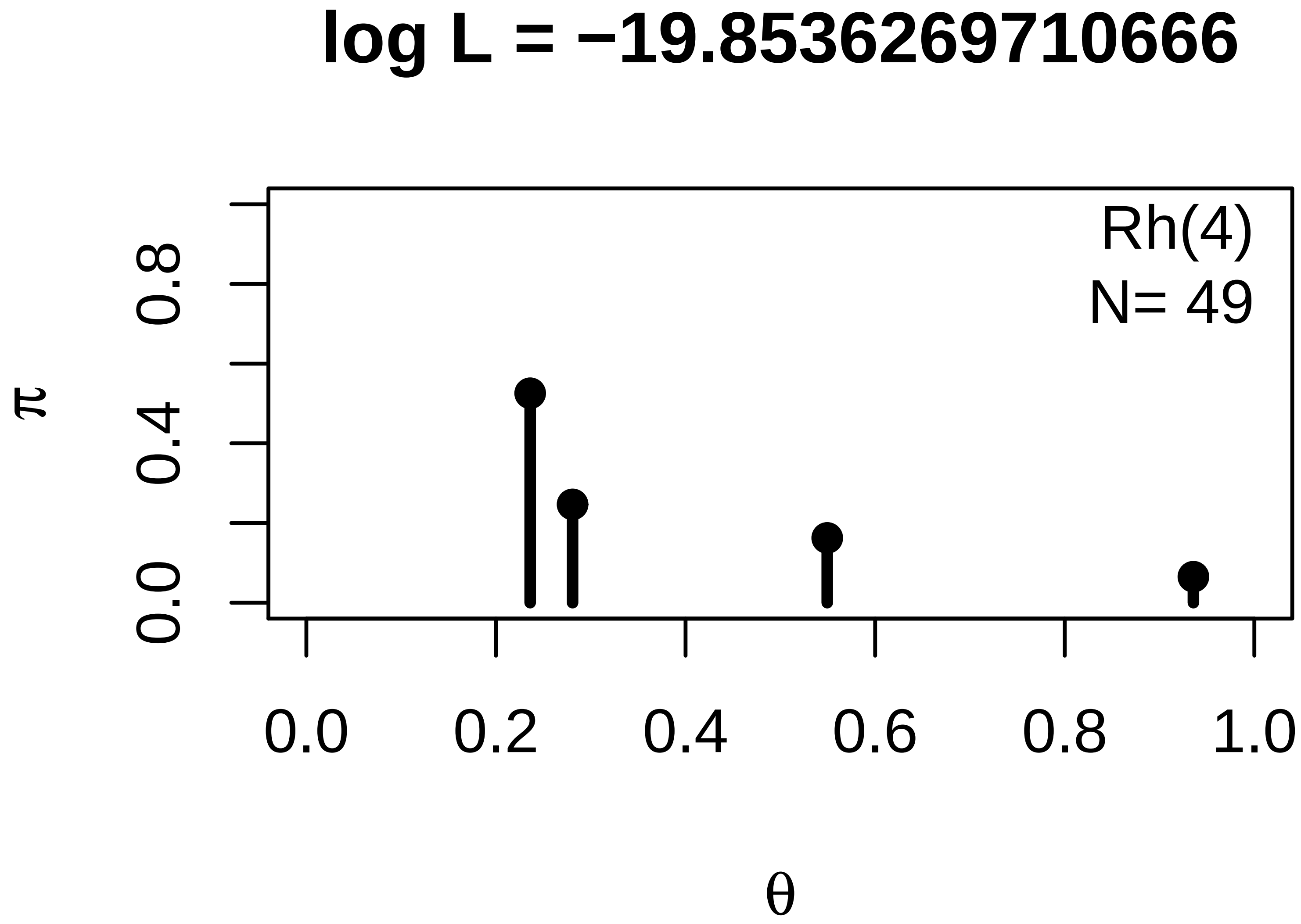
Step 30



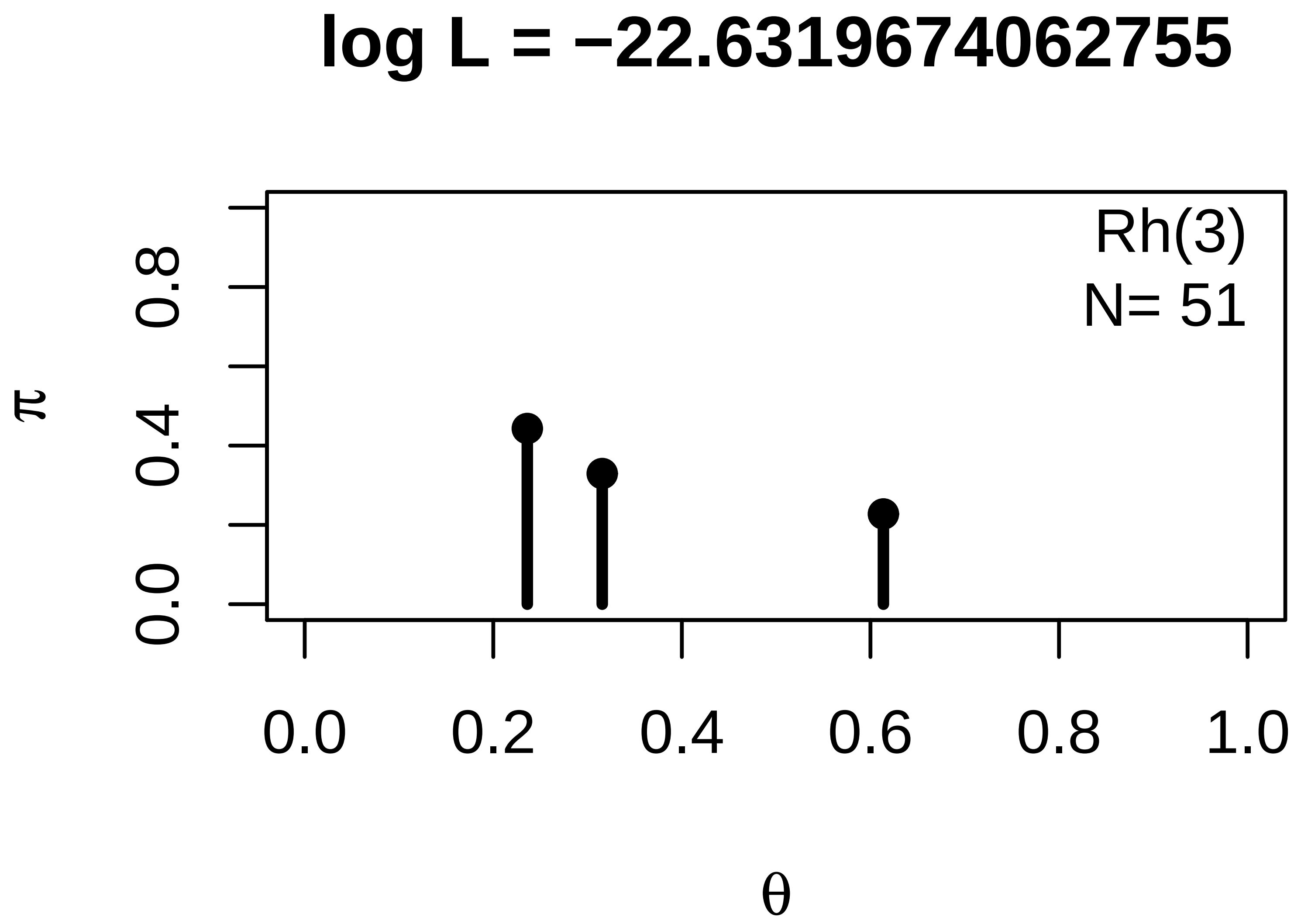
Step 31



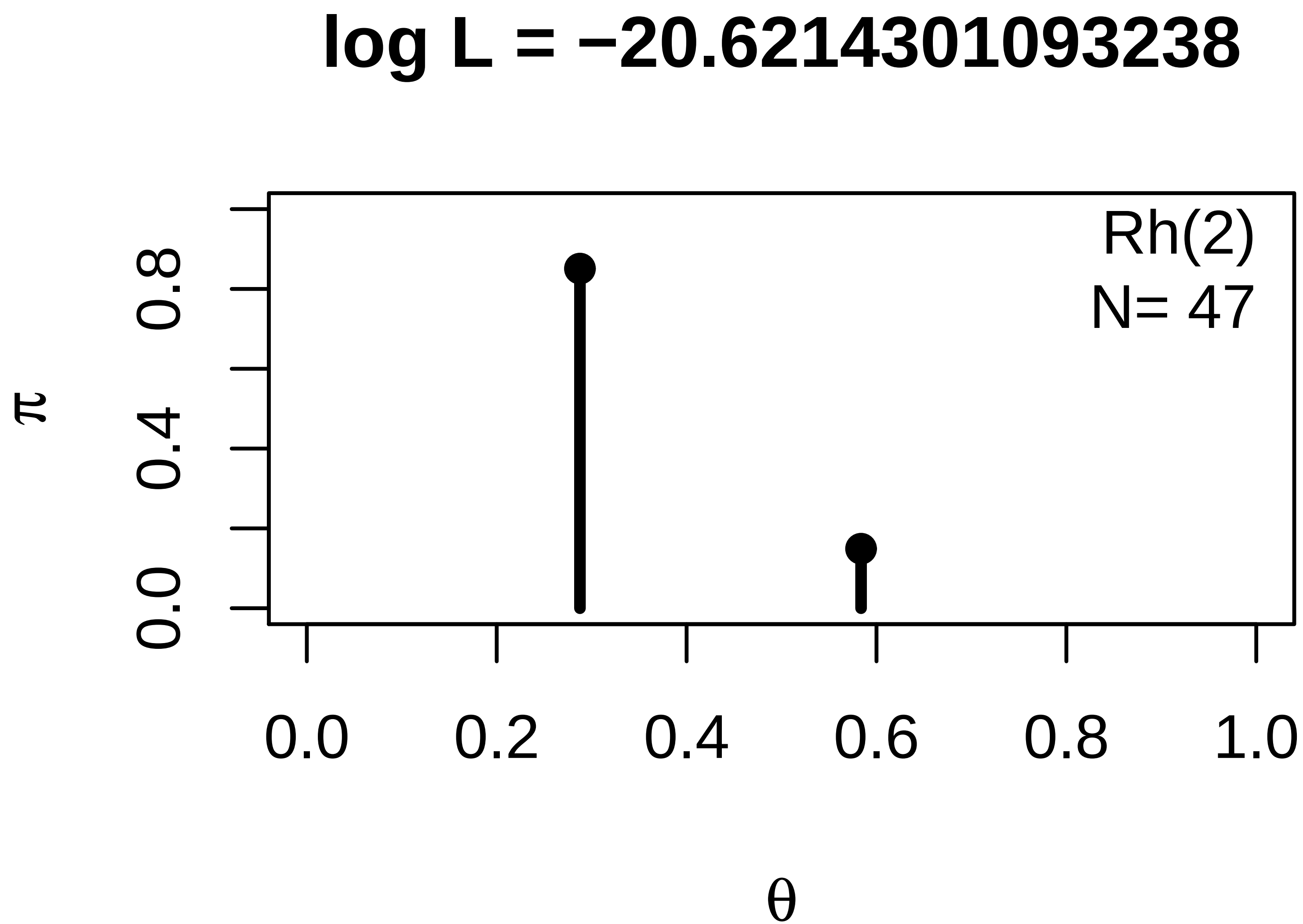
Step 32



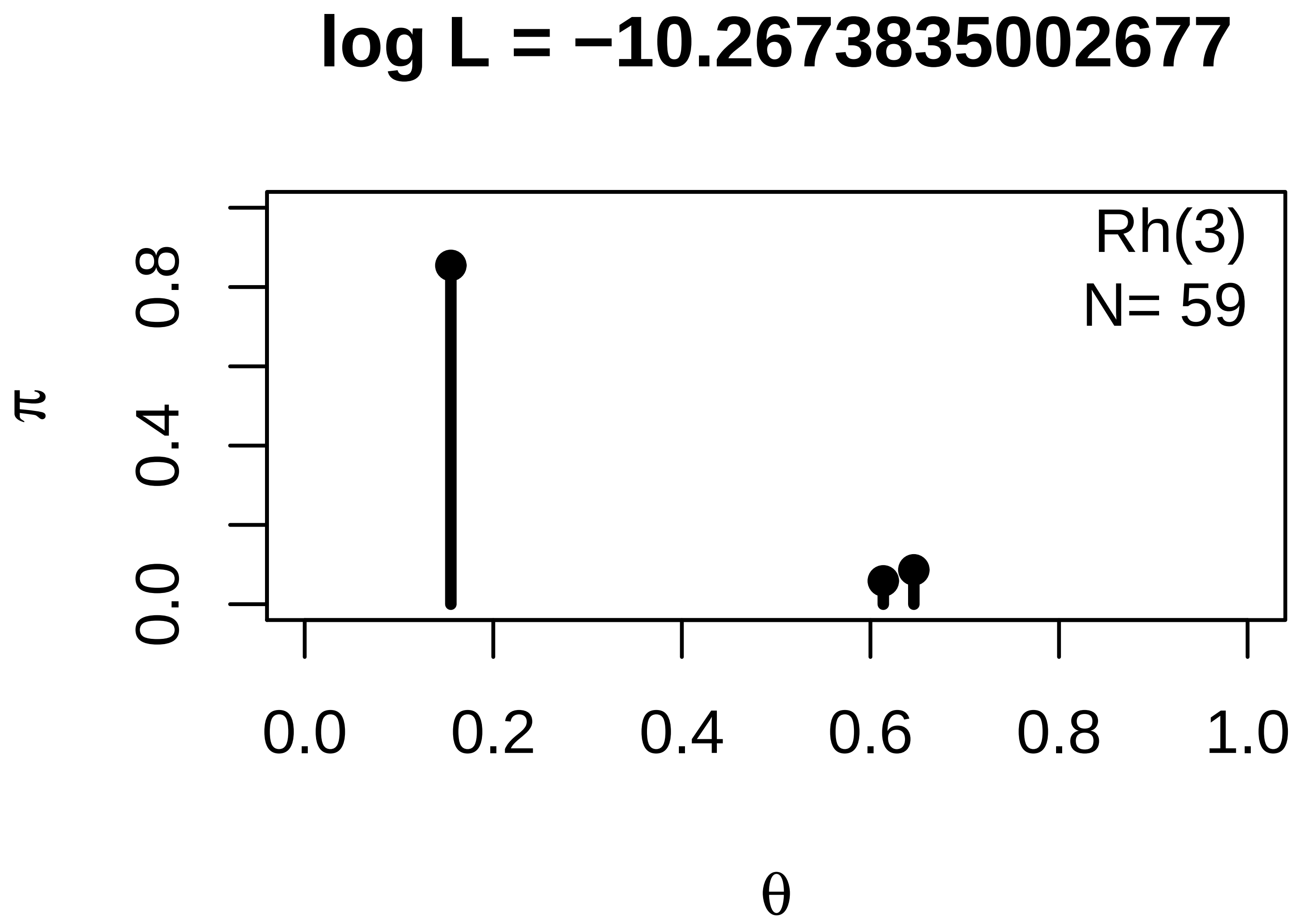
Step 33



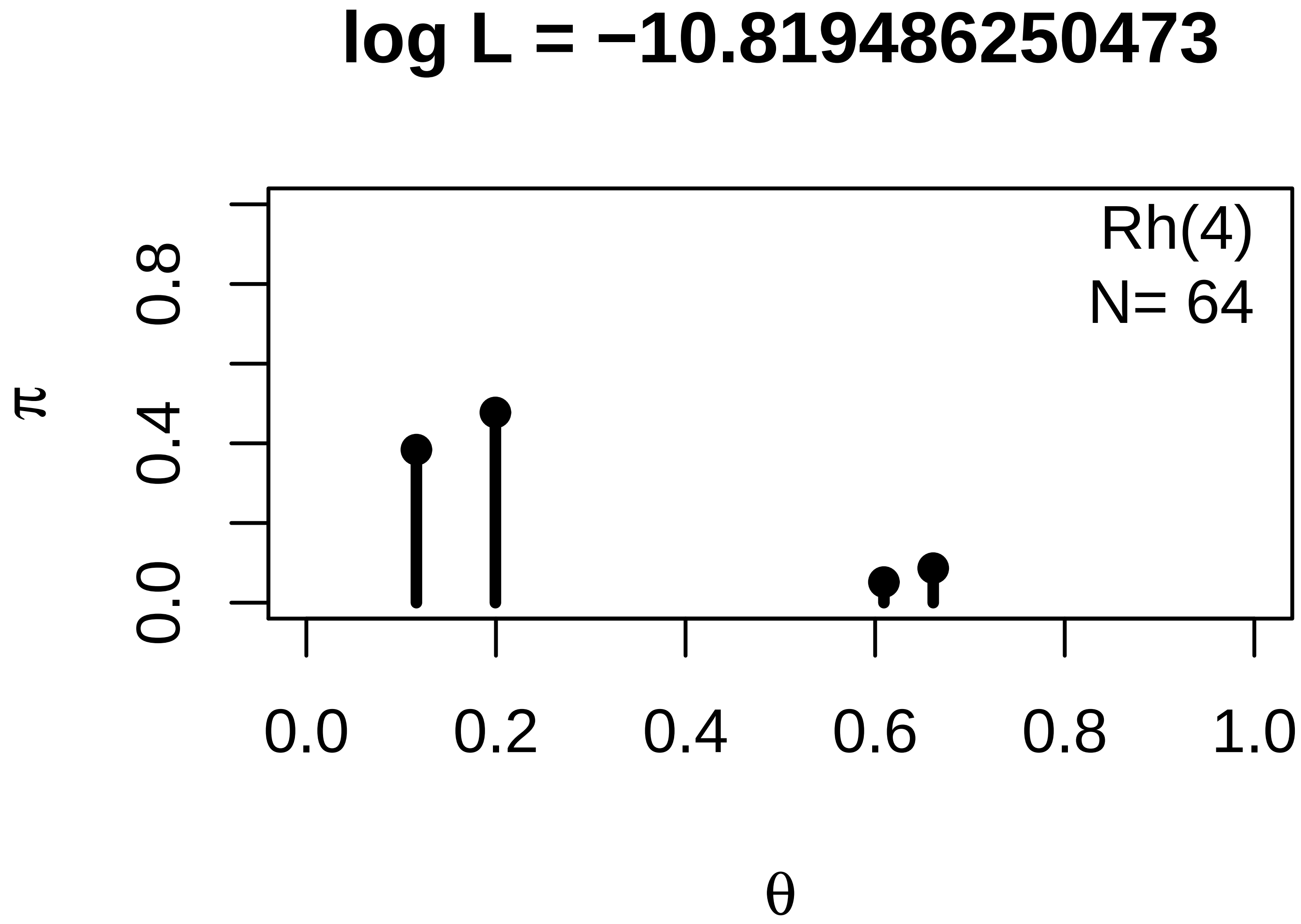
Step 34



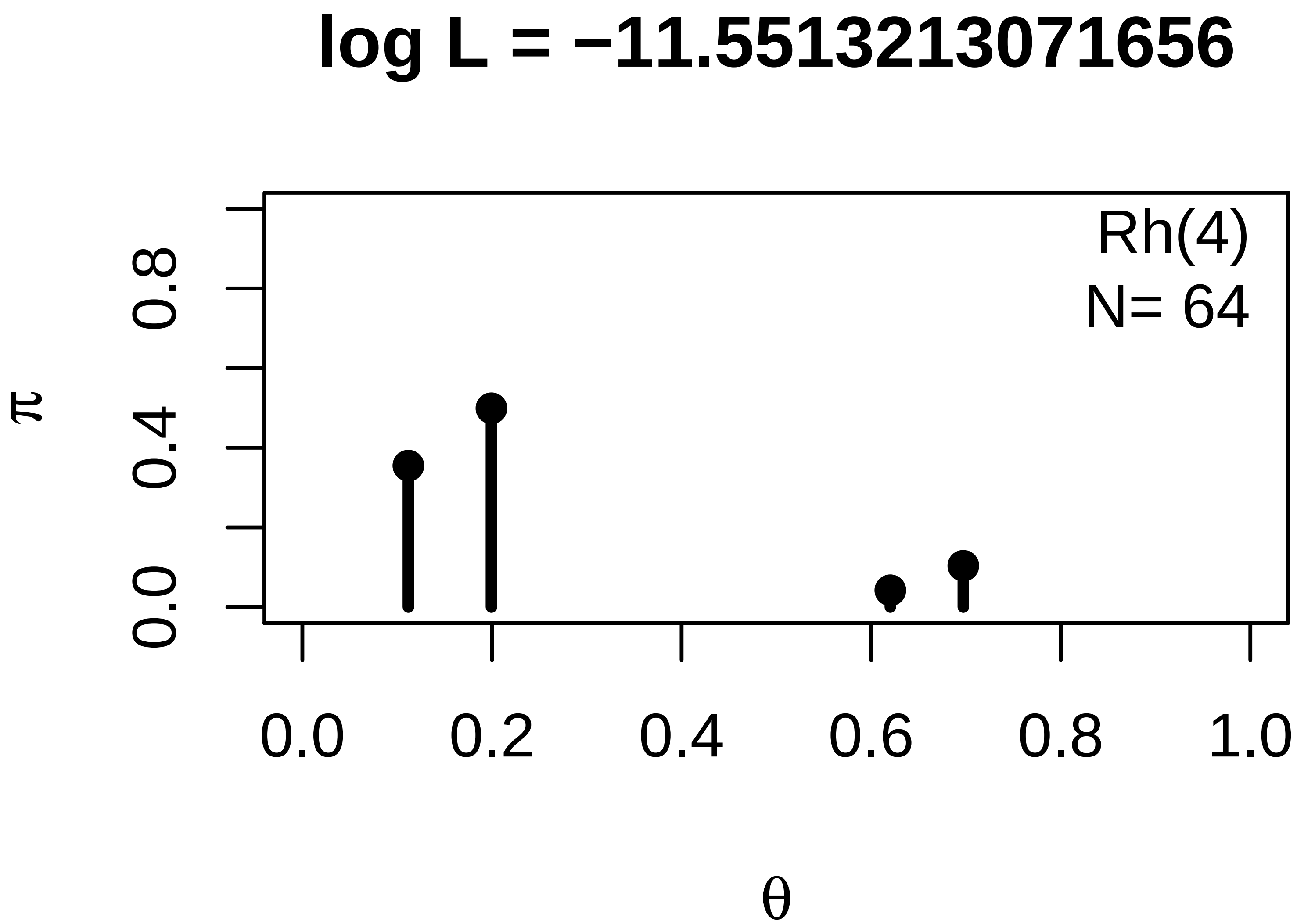
Step 35



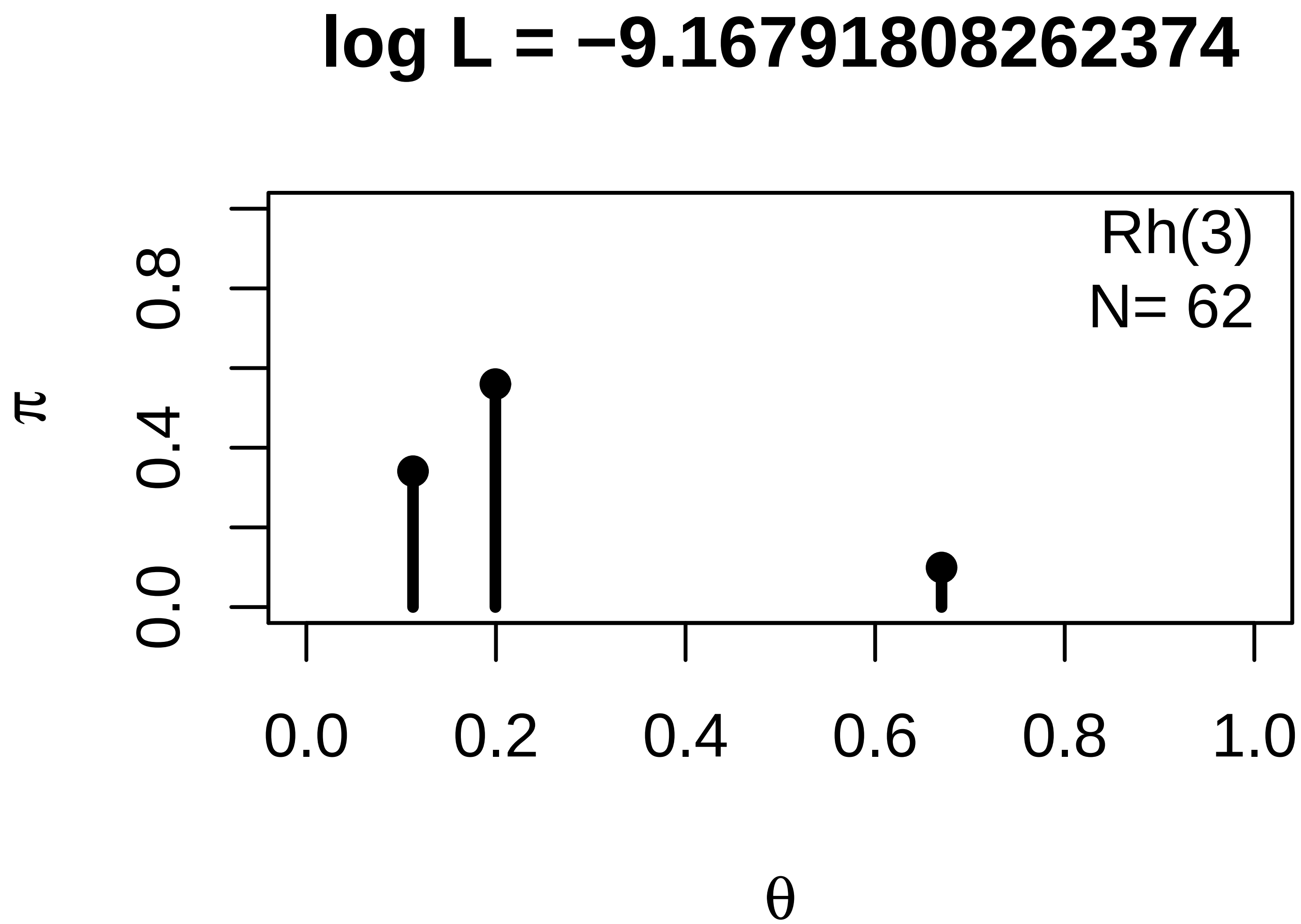
Step 36



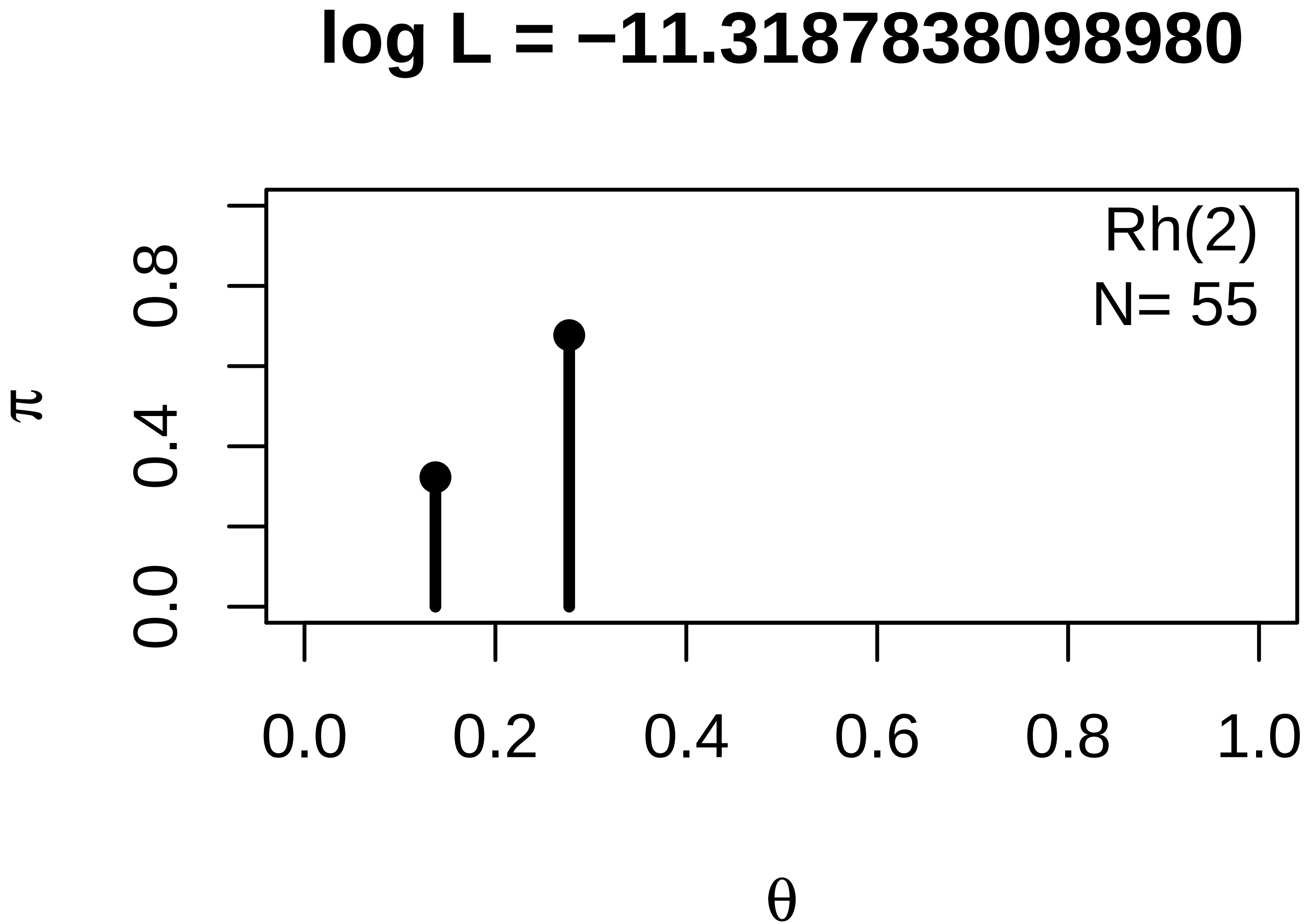
Step 37



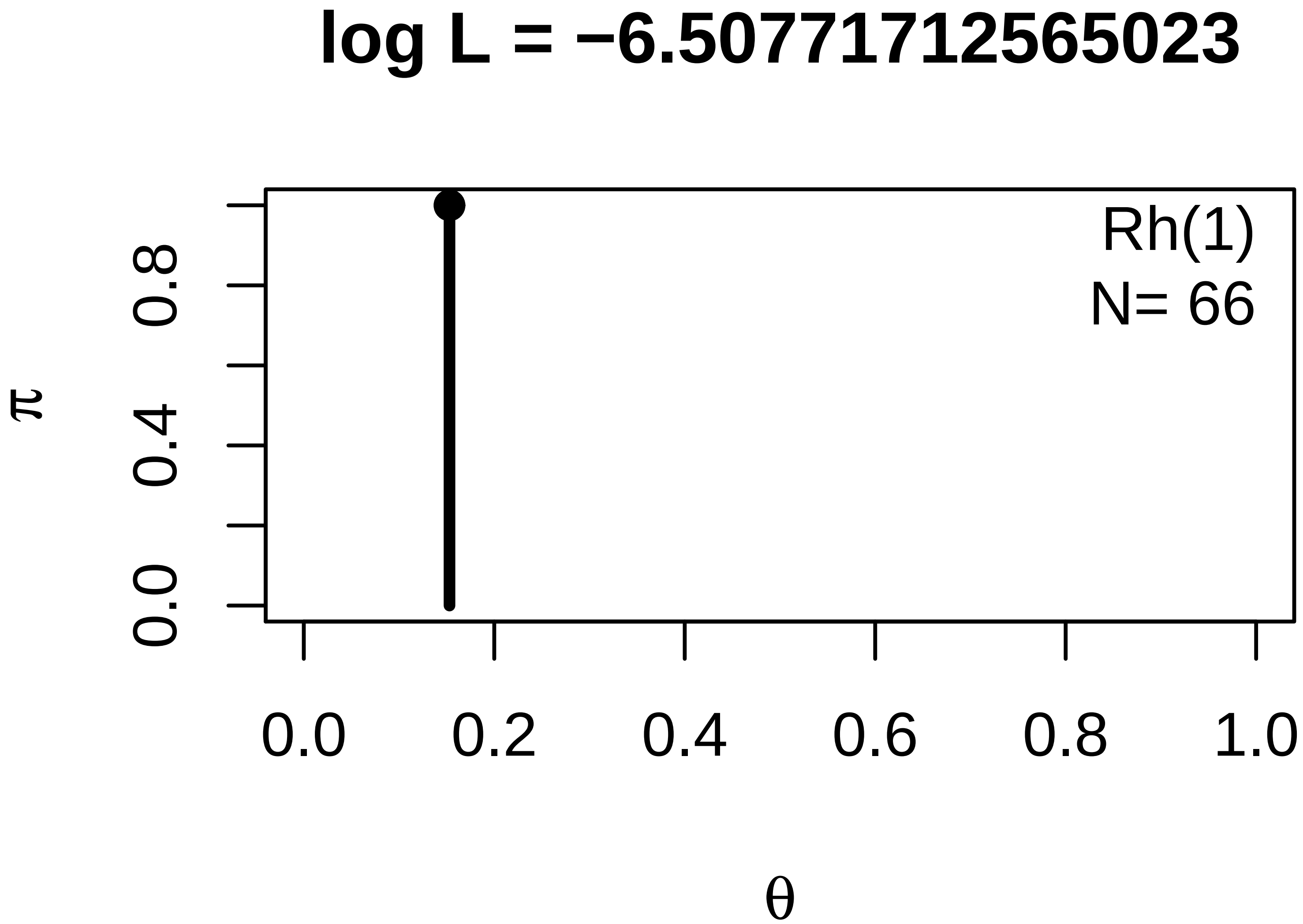
Step 38



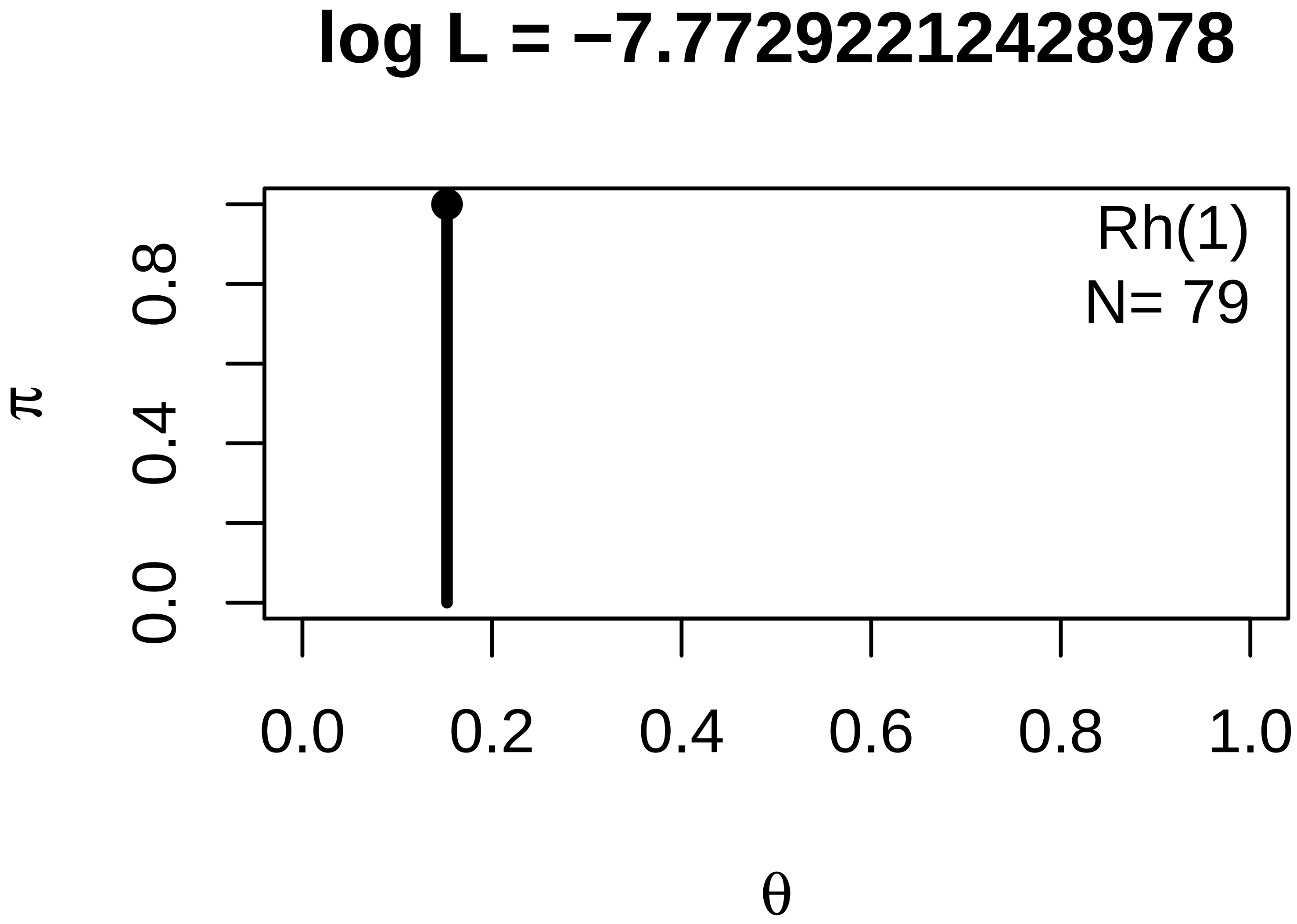
Step 39



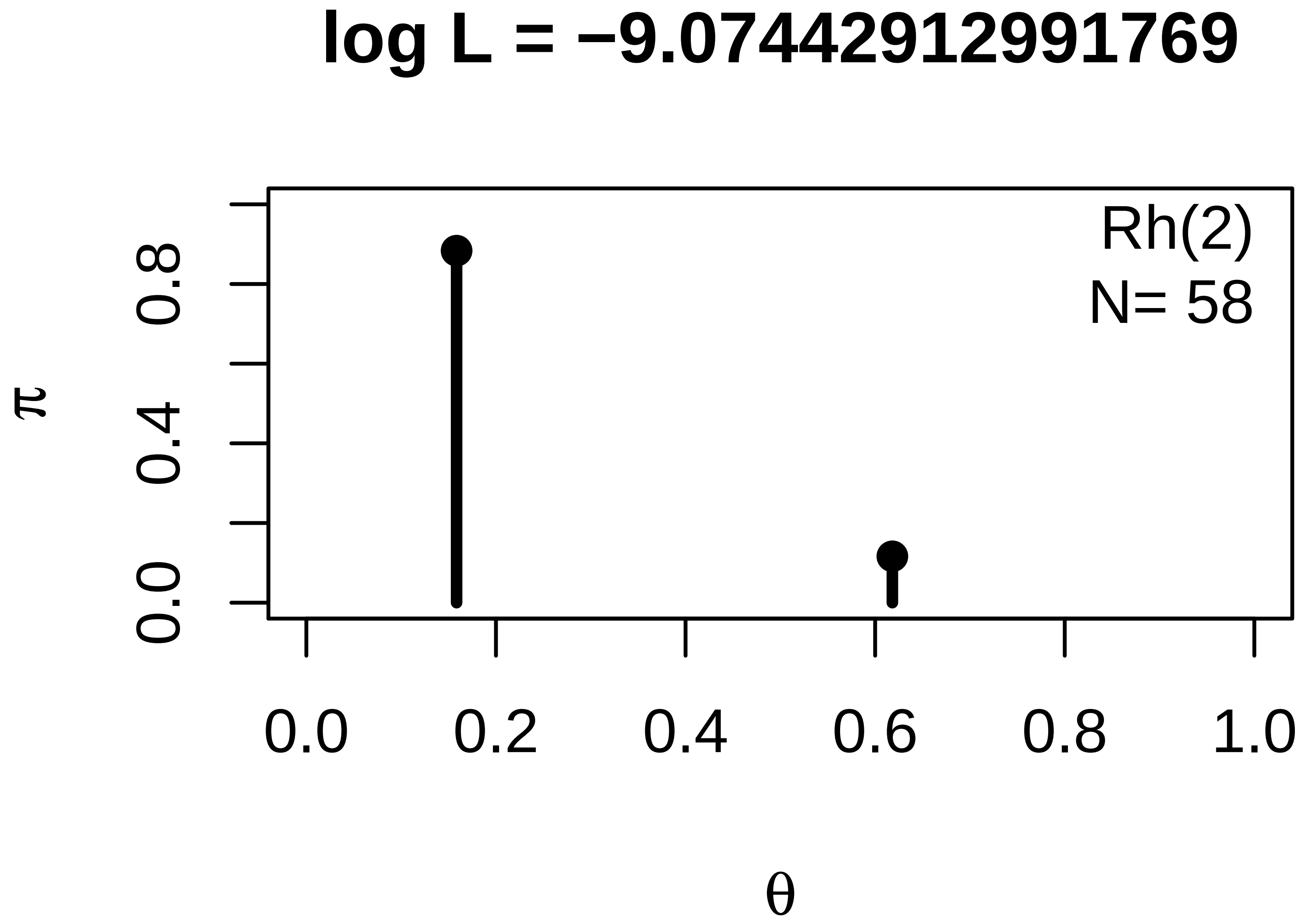
Step 40



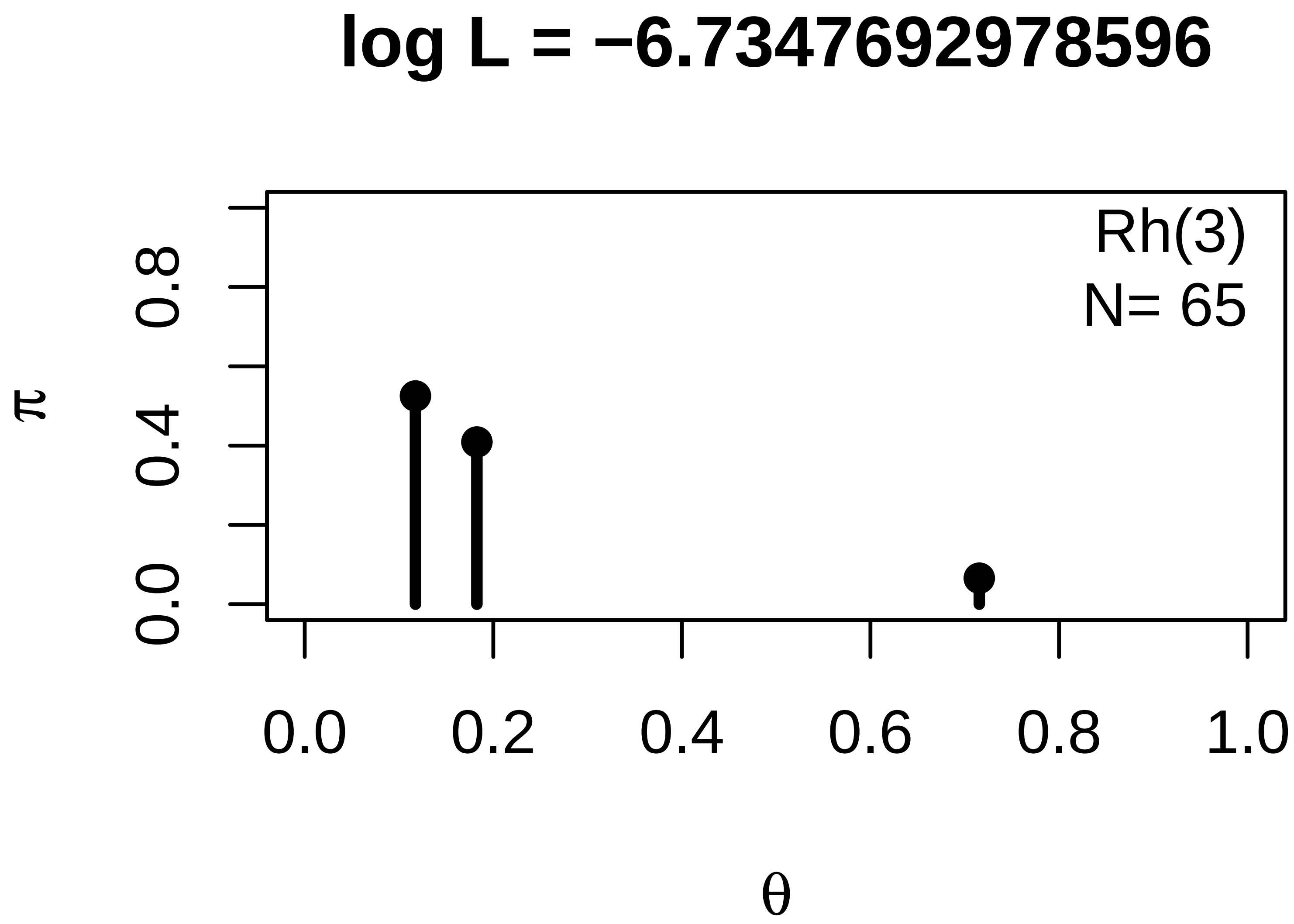
Step 41



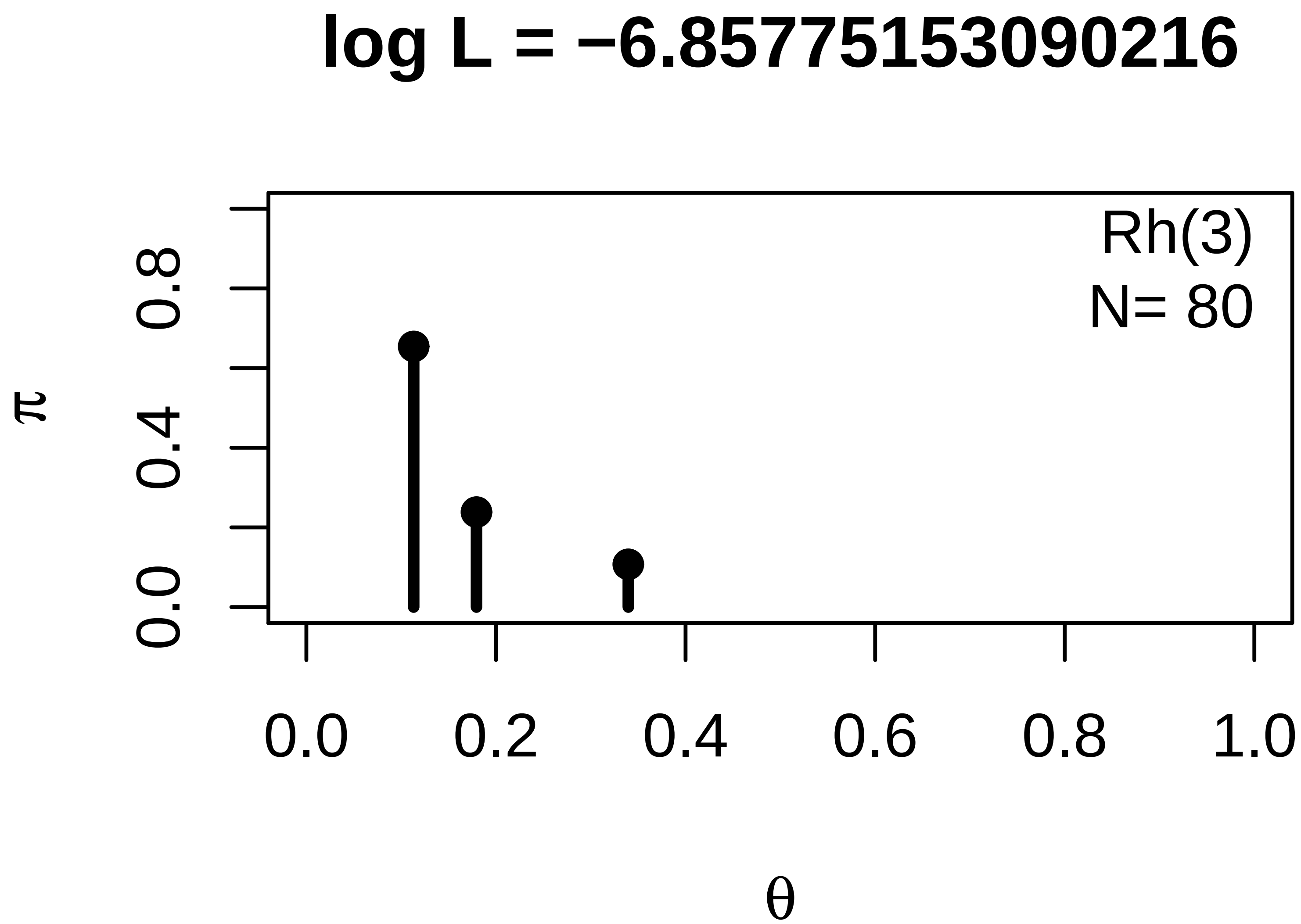
Step 42



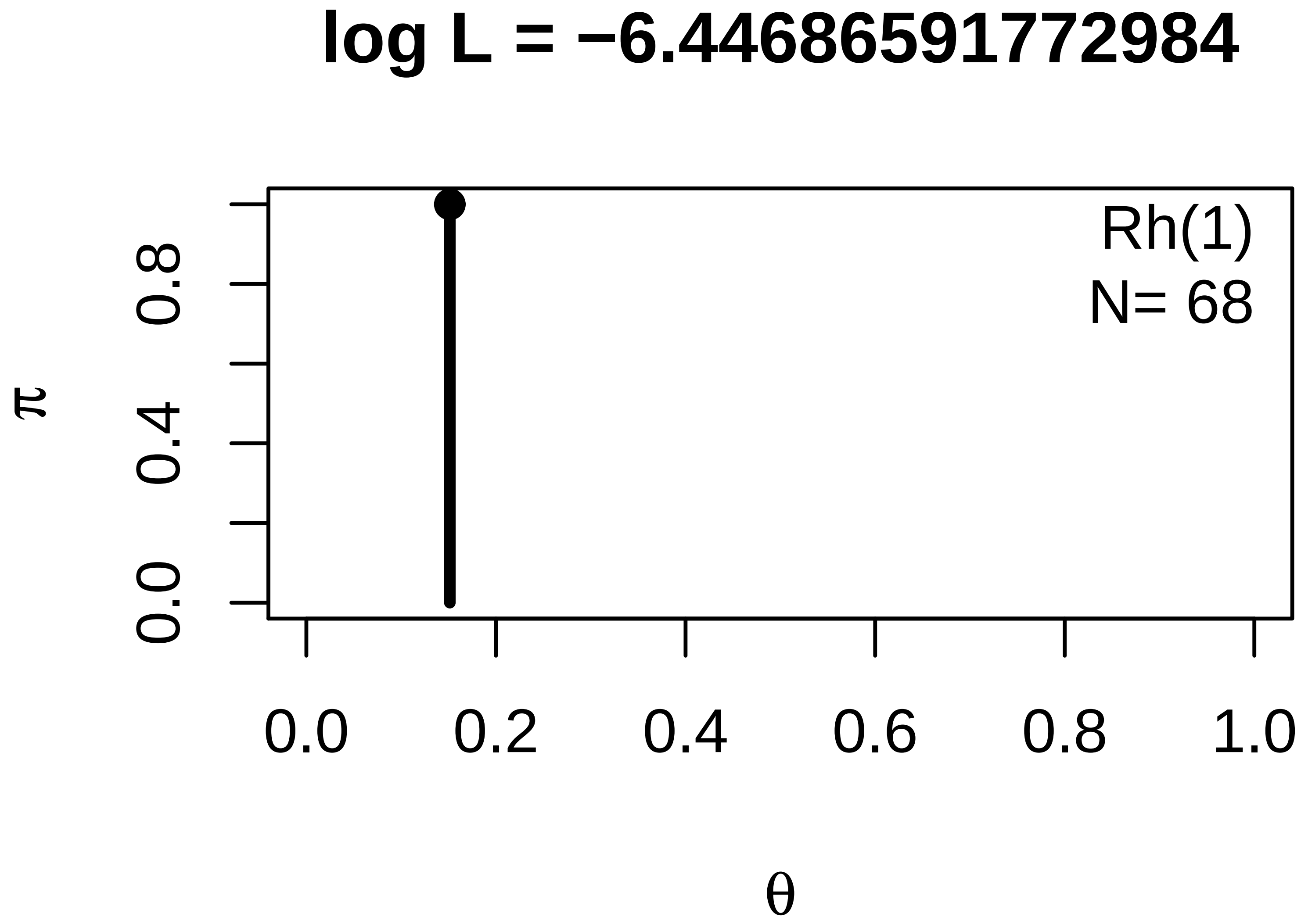
Step 43



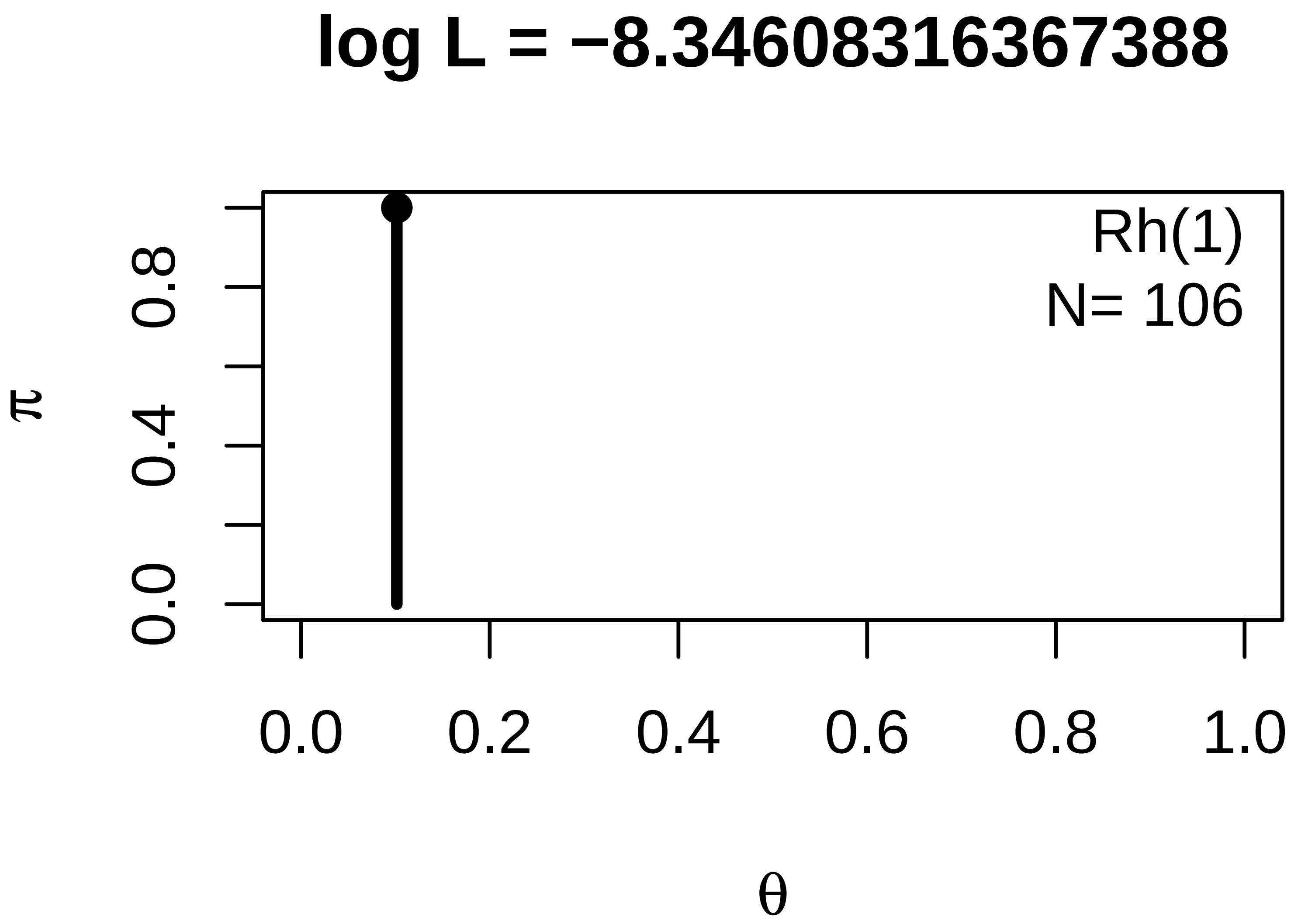
Step 44



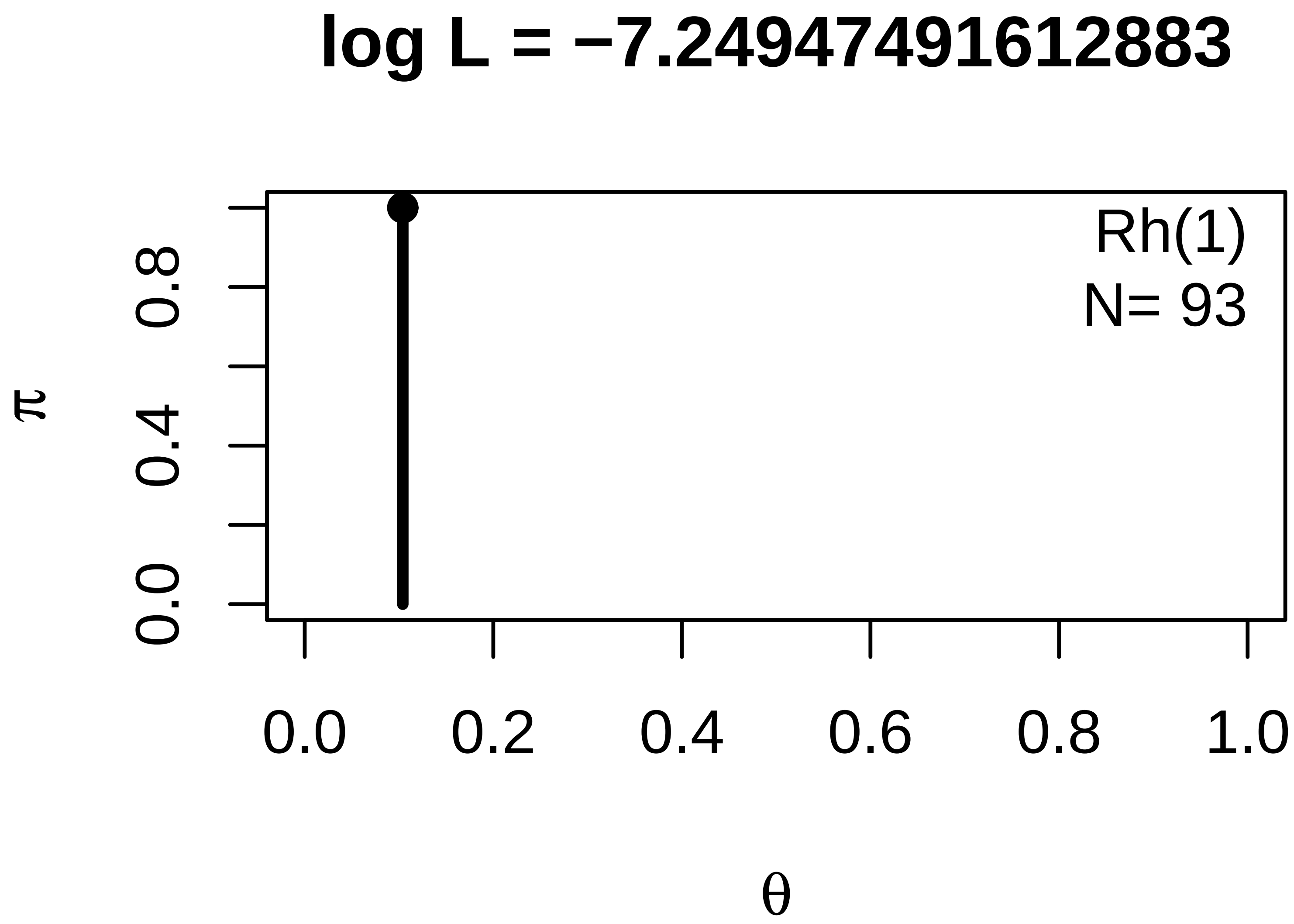
Step 45



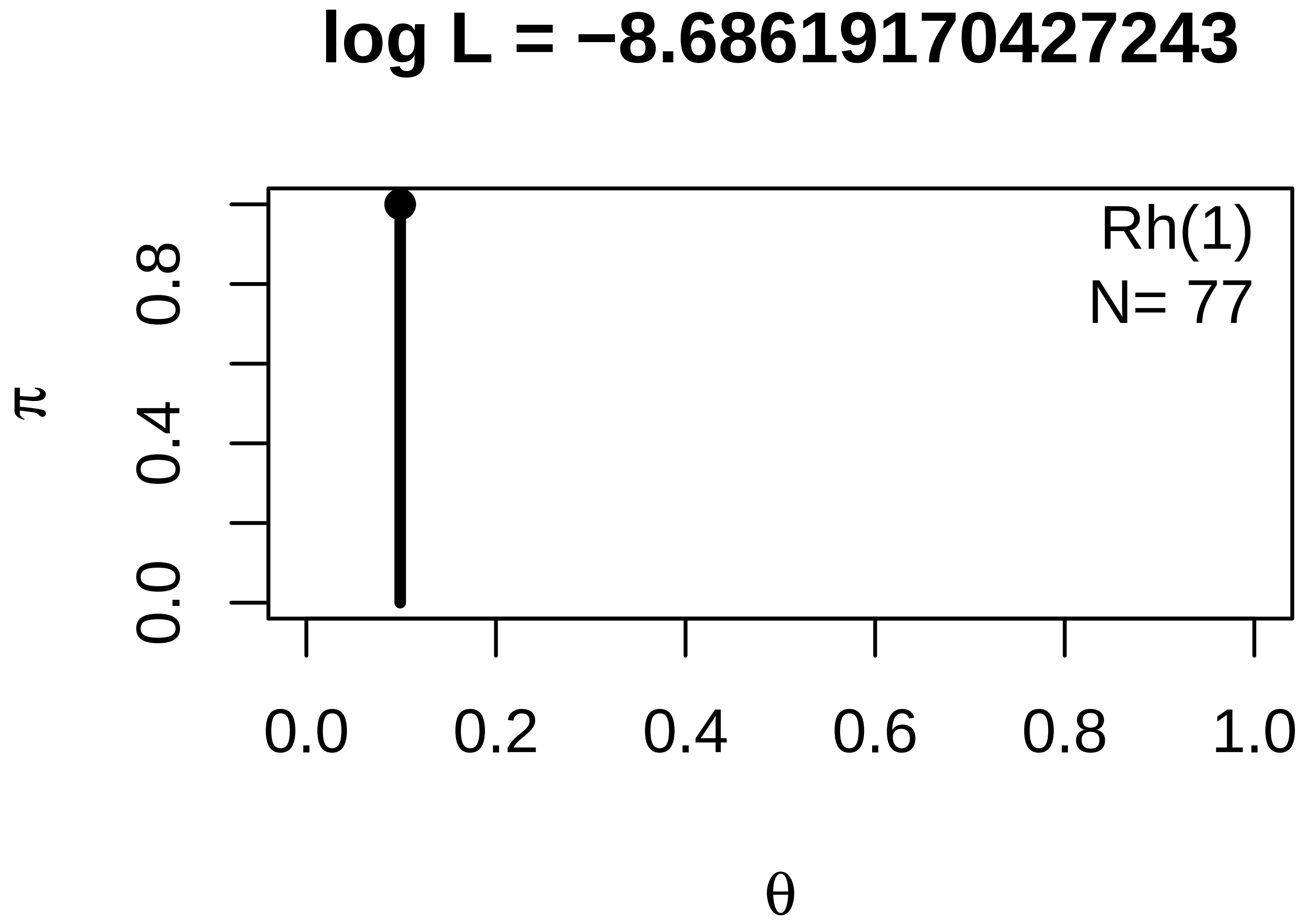
Step 46



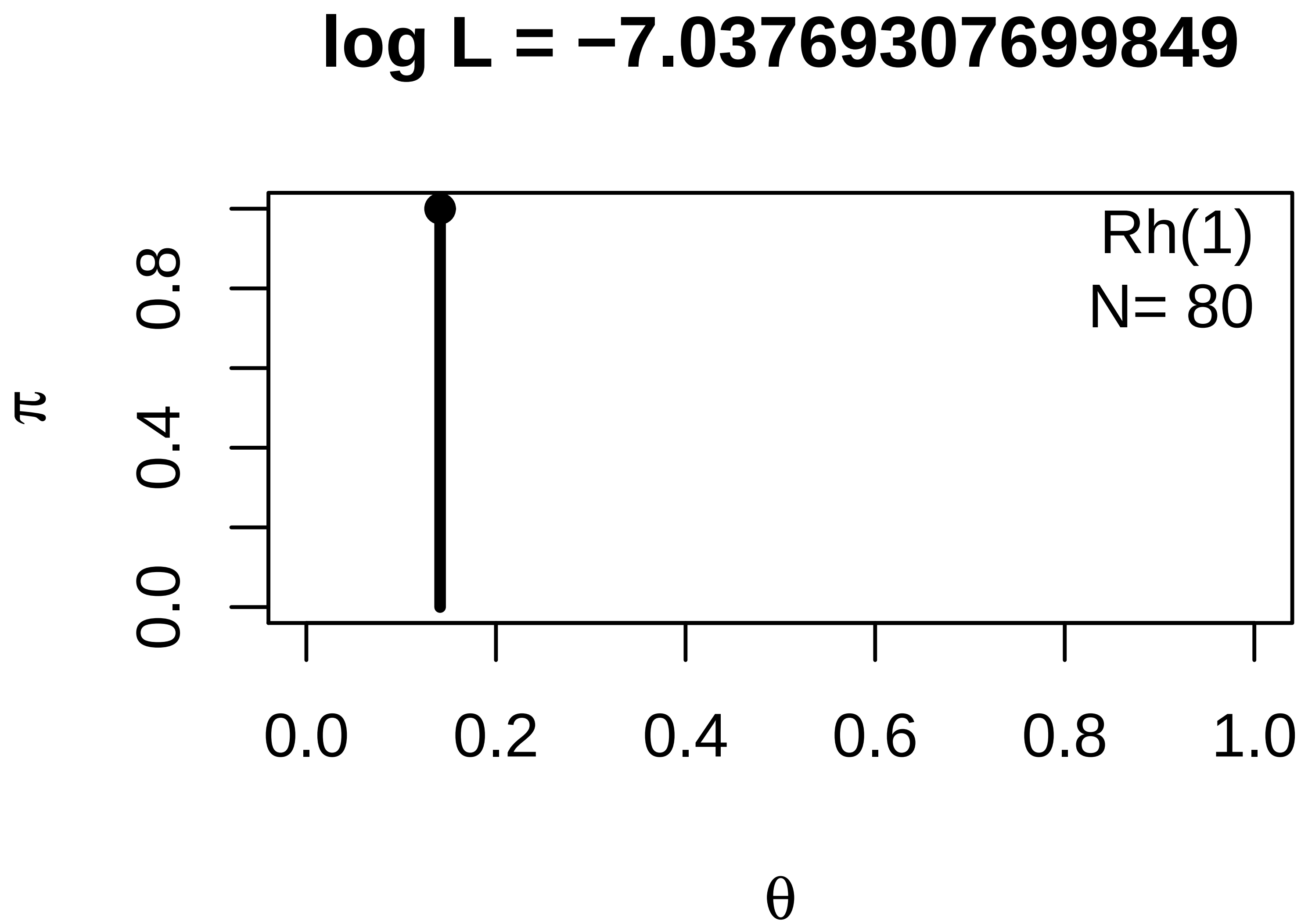
Step 47



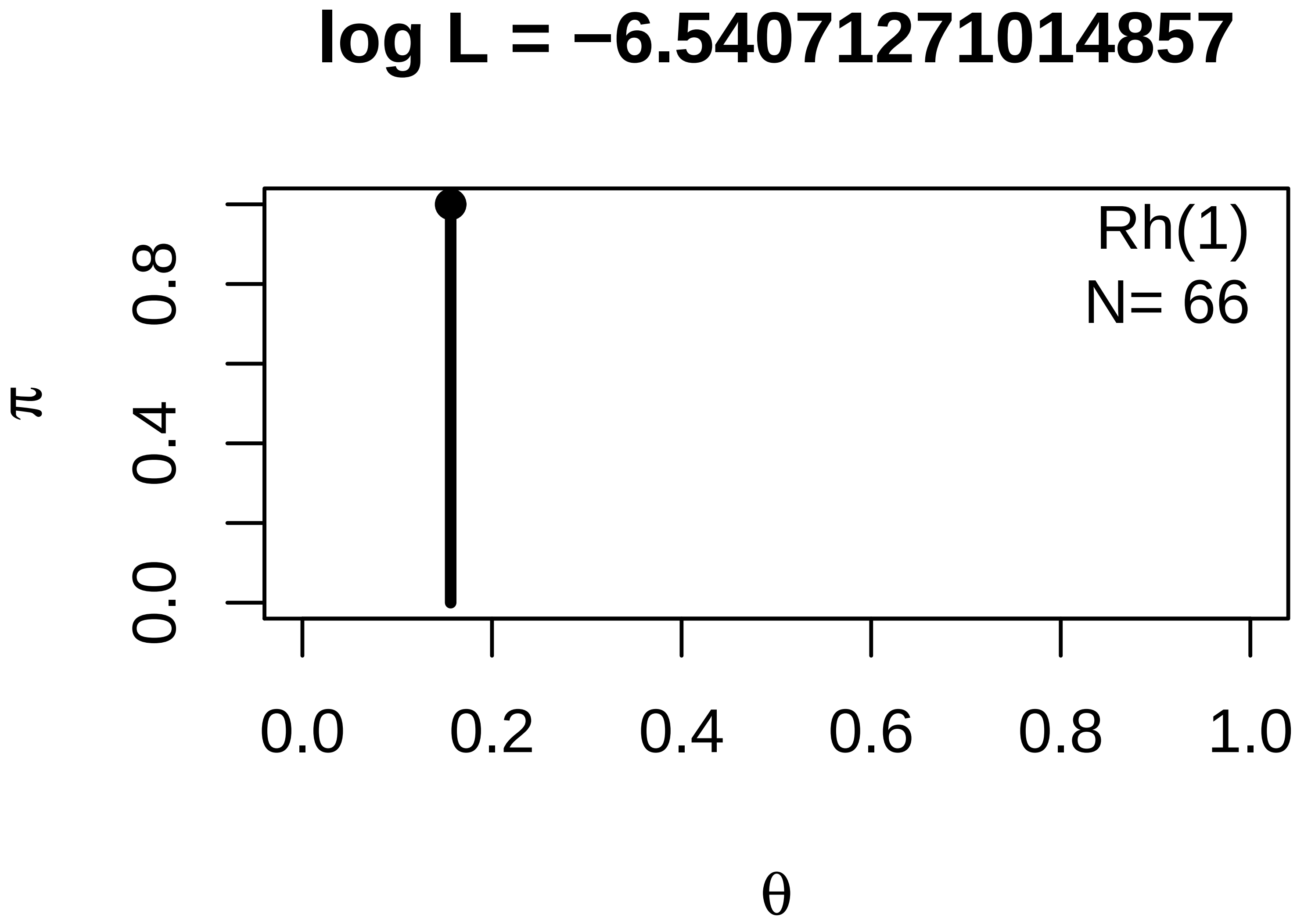
Step 48



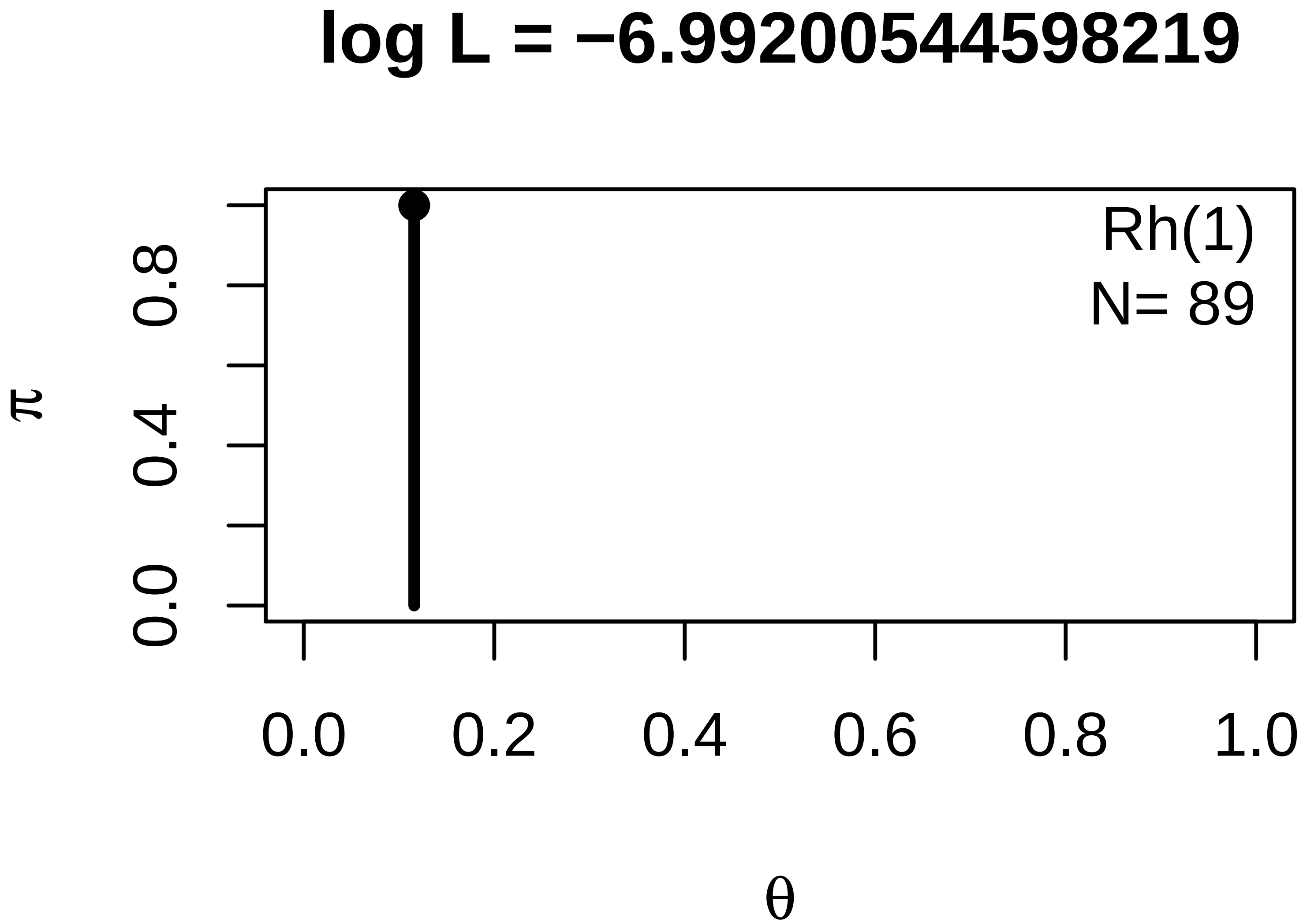
Step 49



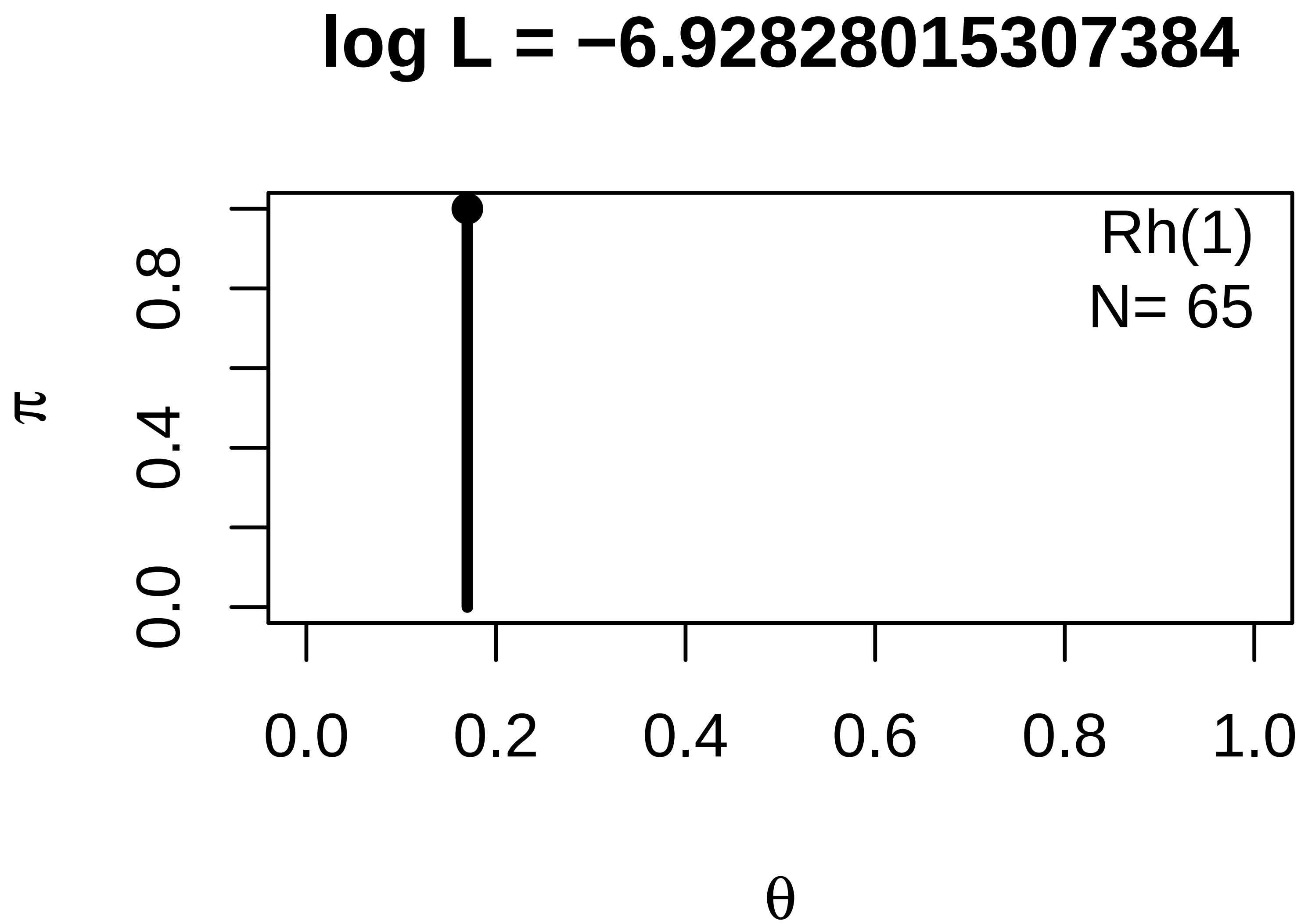
Step 50



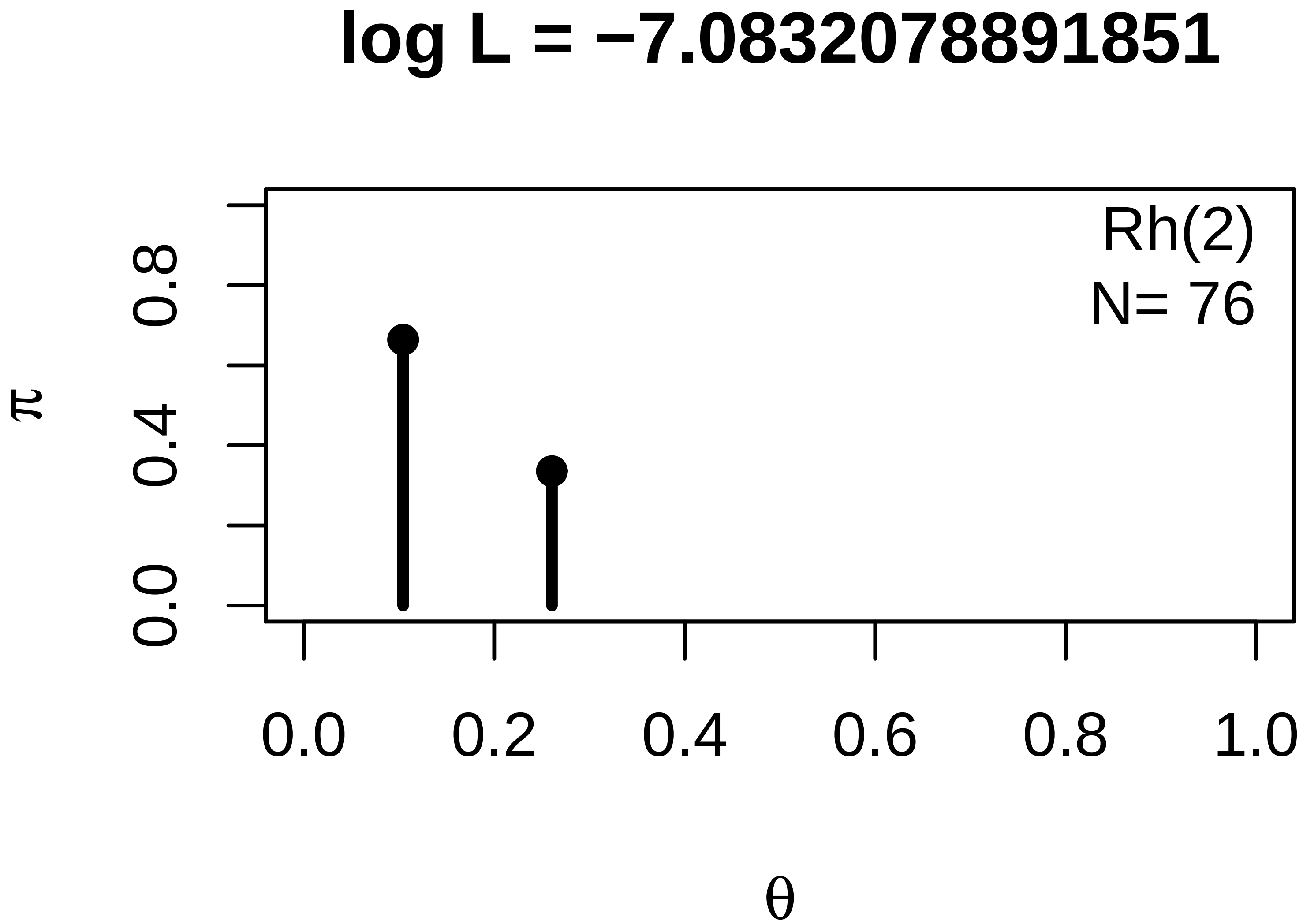
Step 51



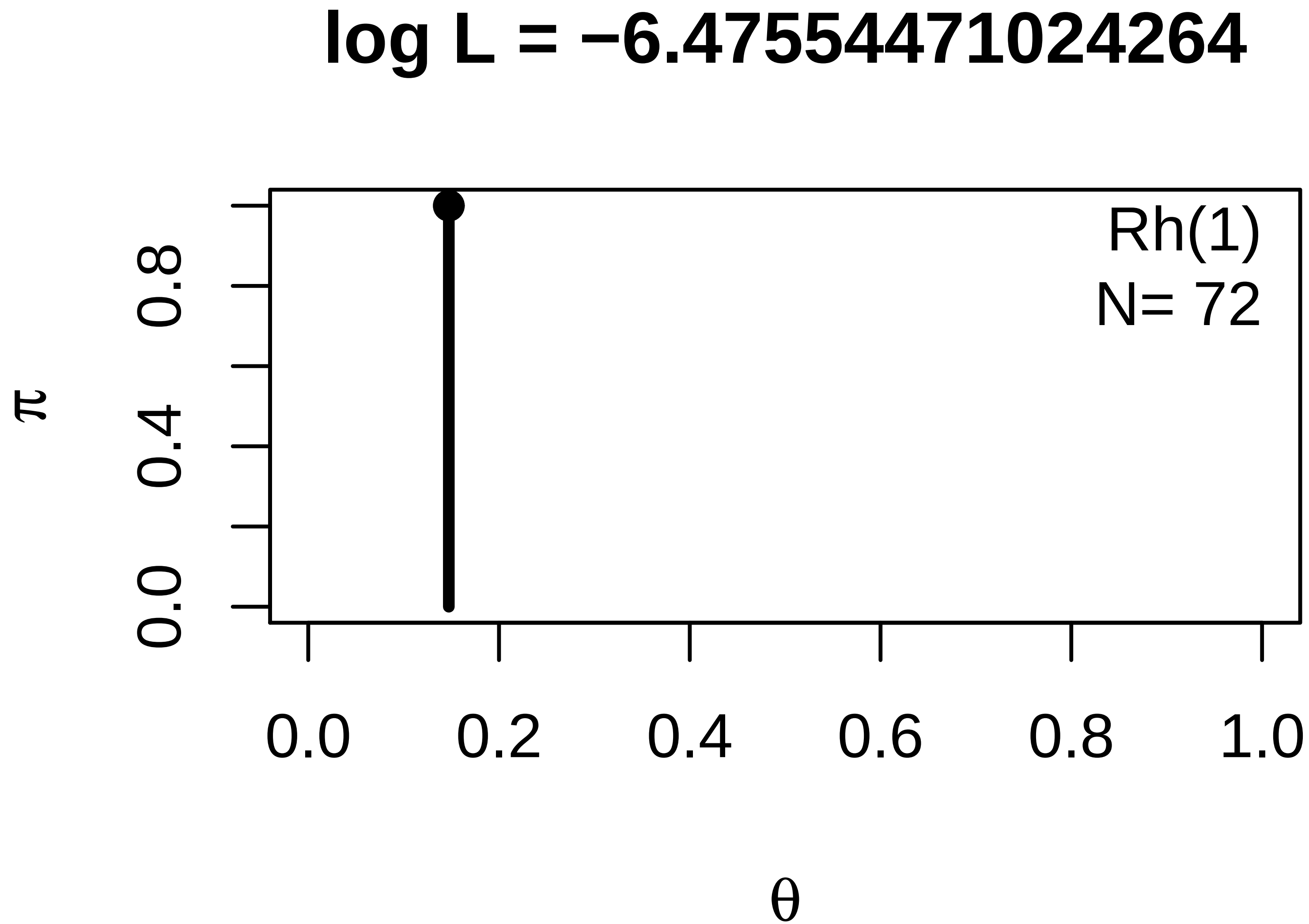
Step 52



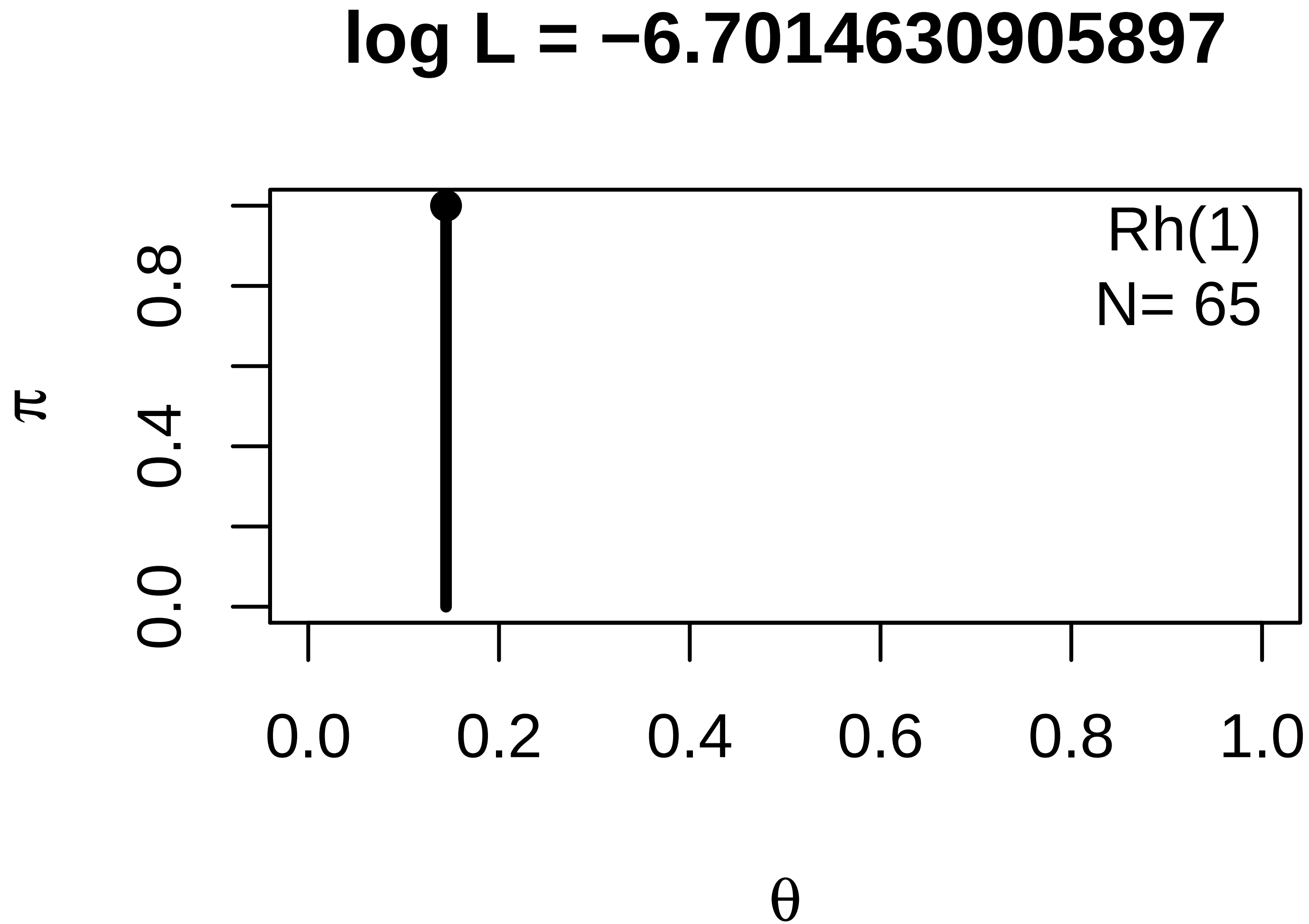
Step 53



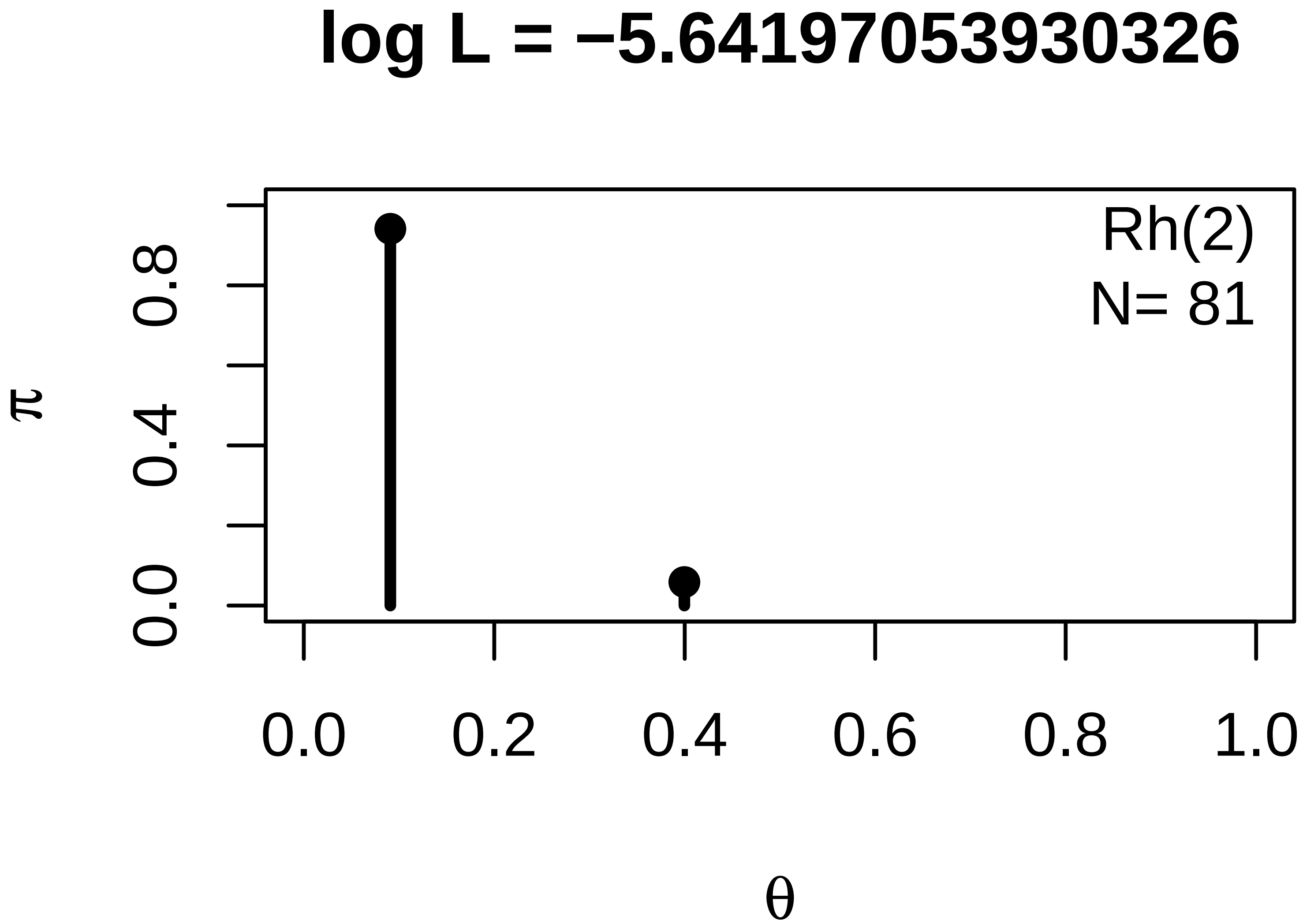
Step 54



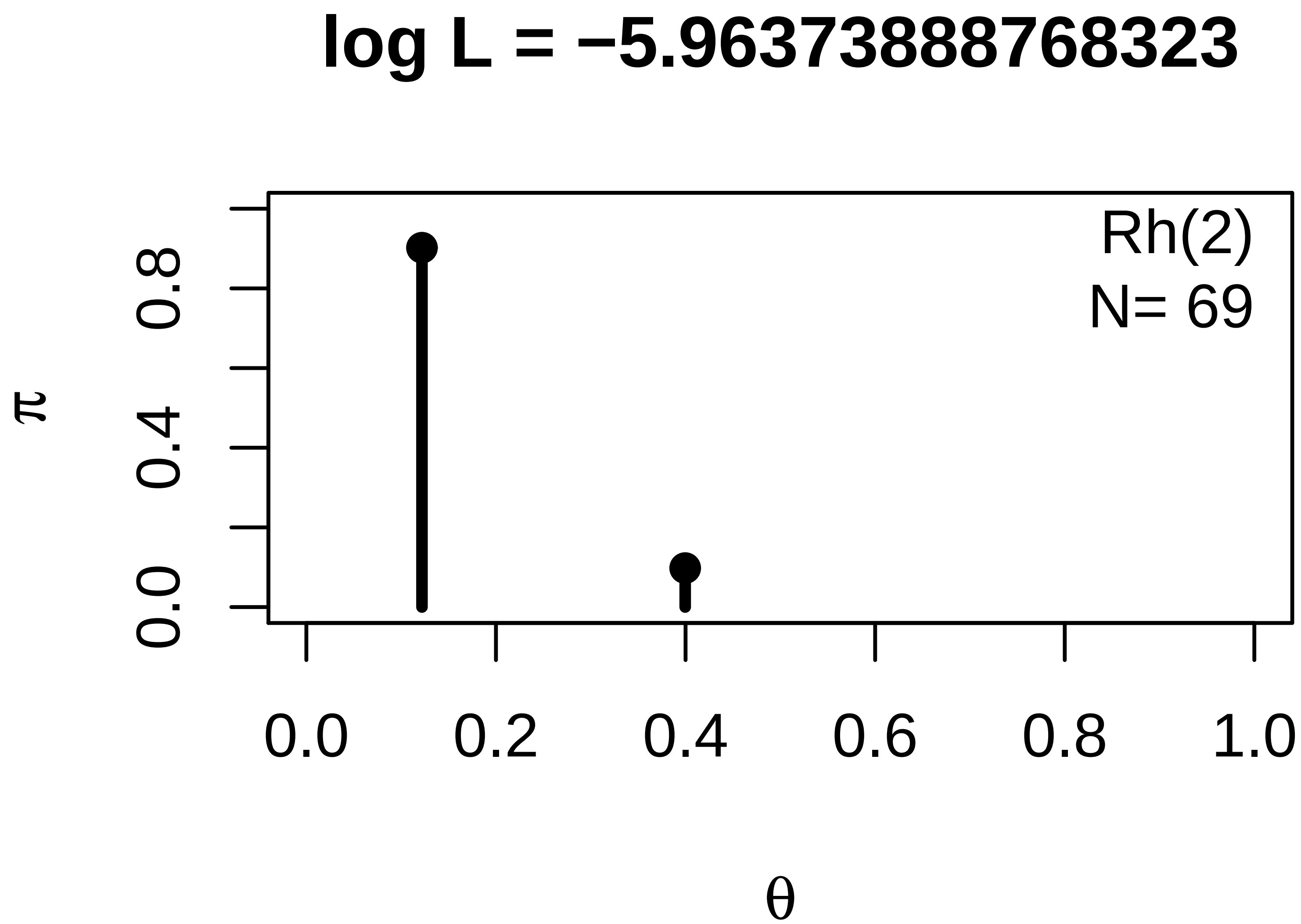
Step 55



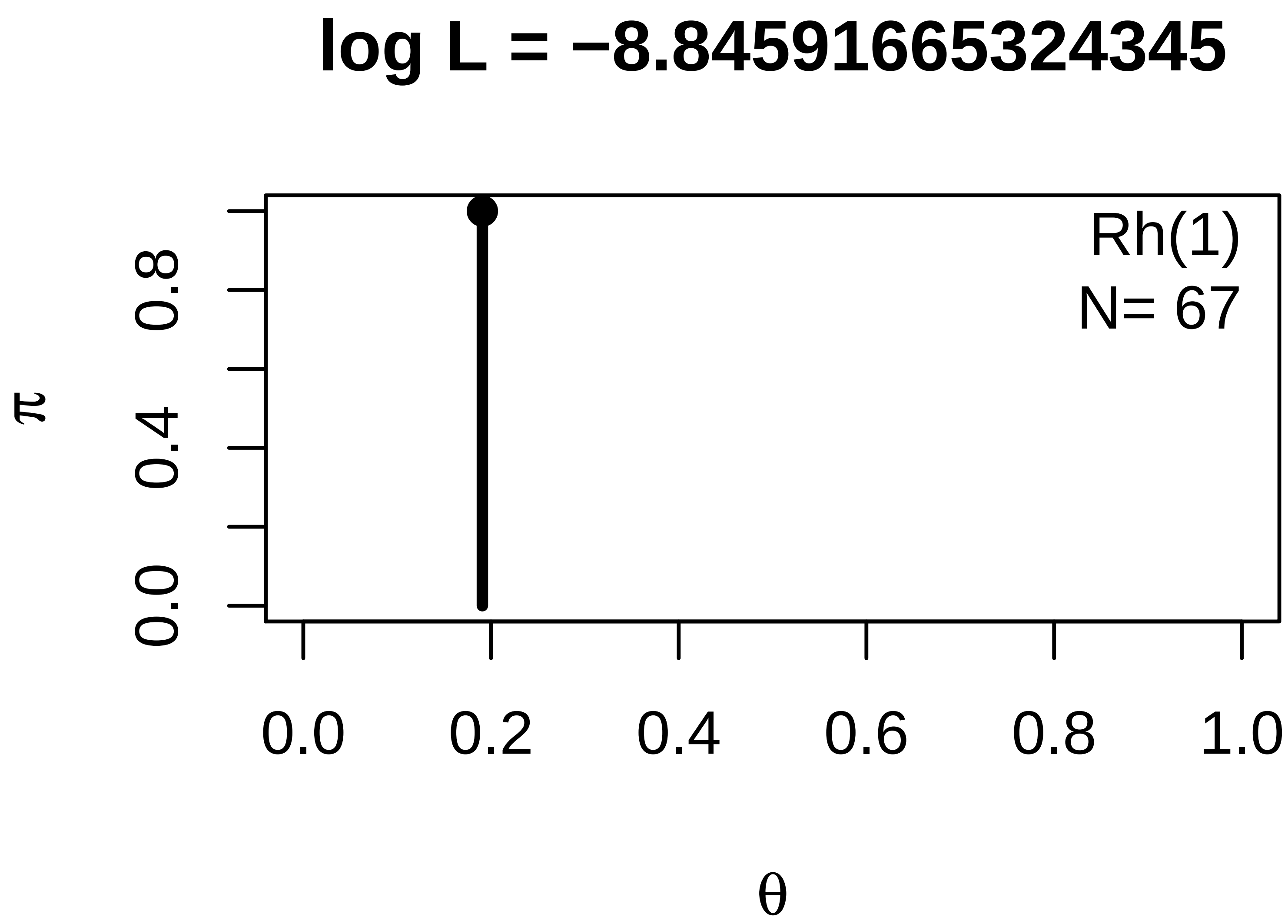
Step 56



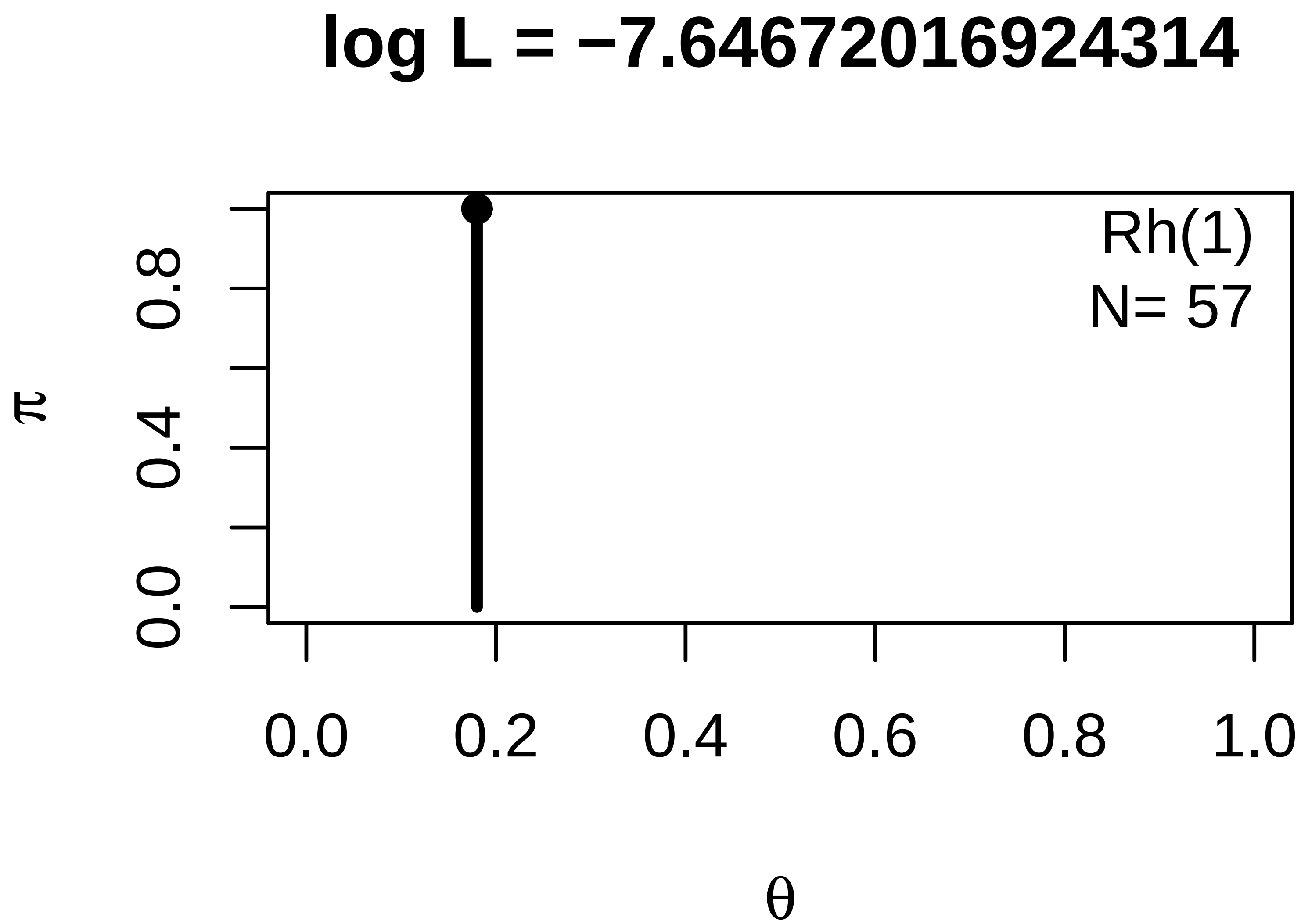
Step 57



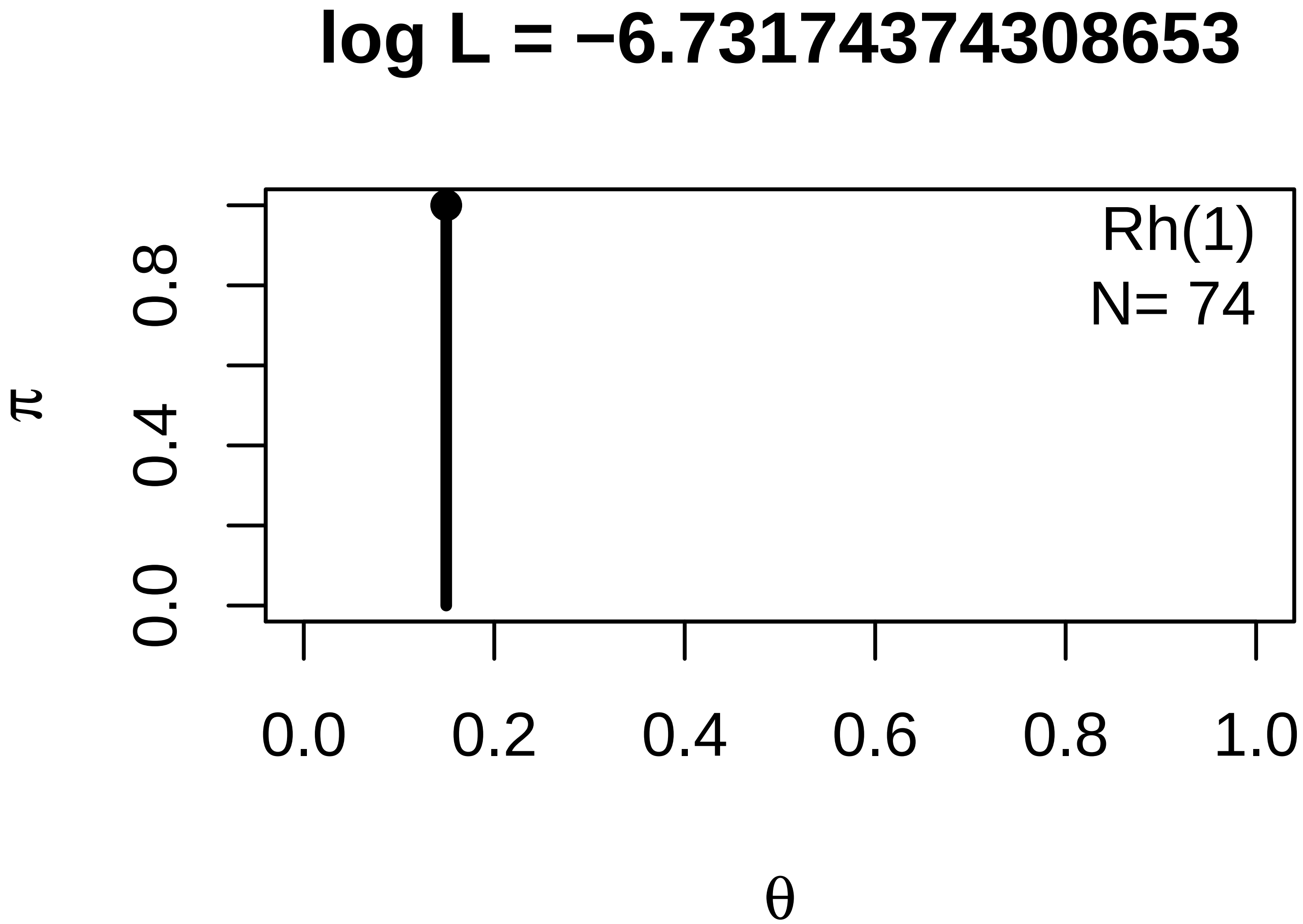
Step 58



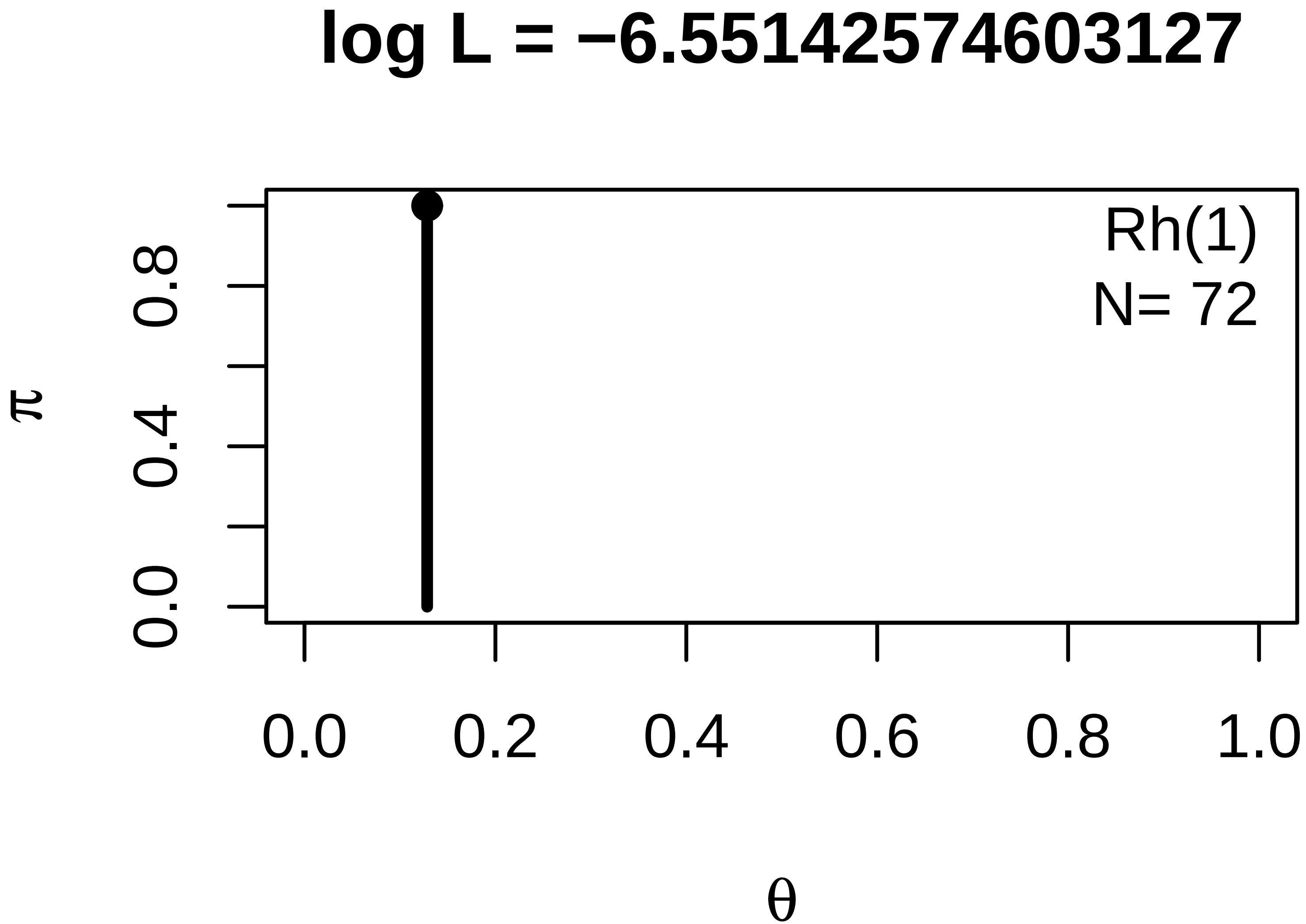
Step 59



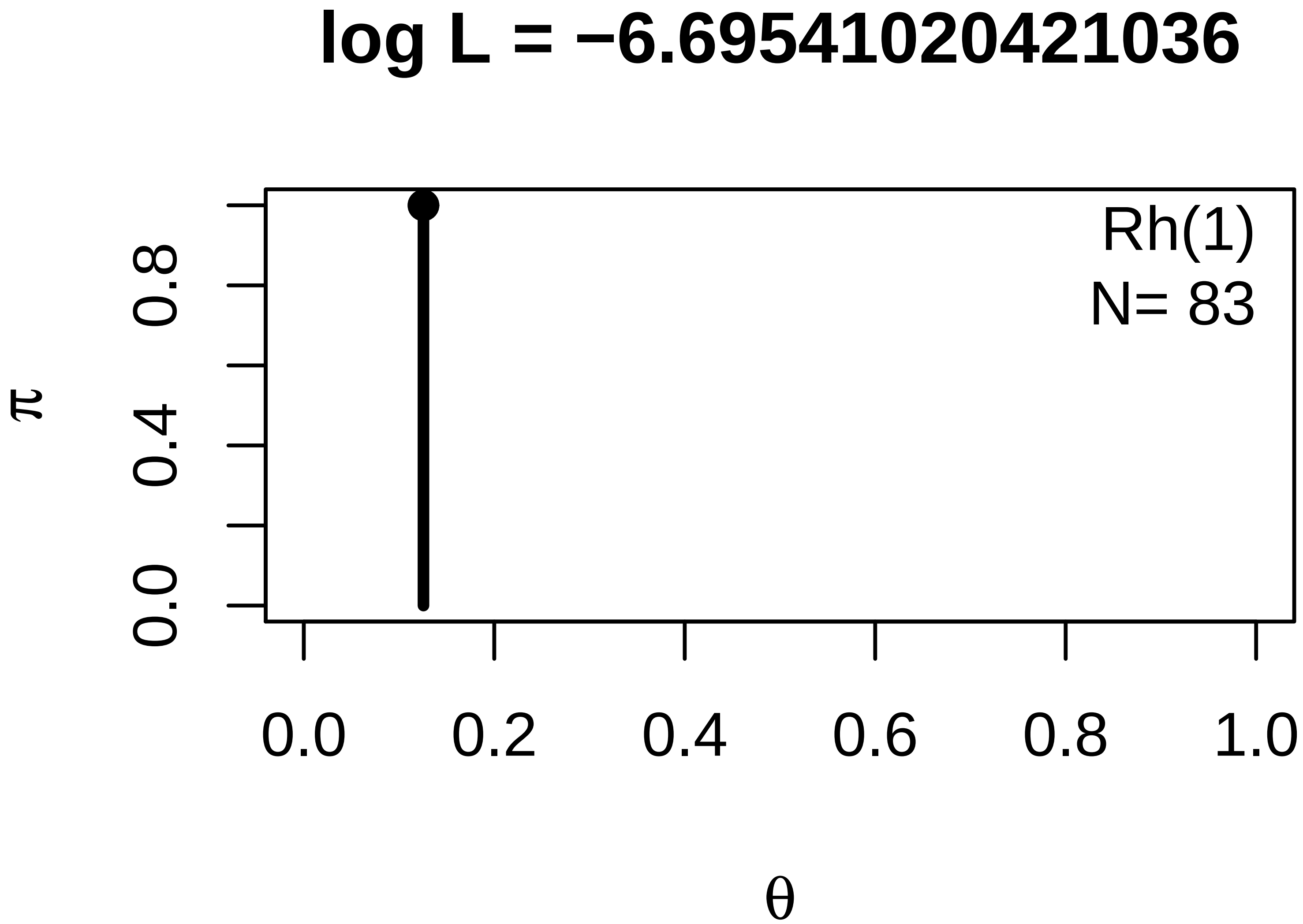
Step 60



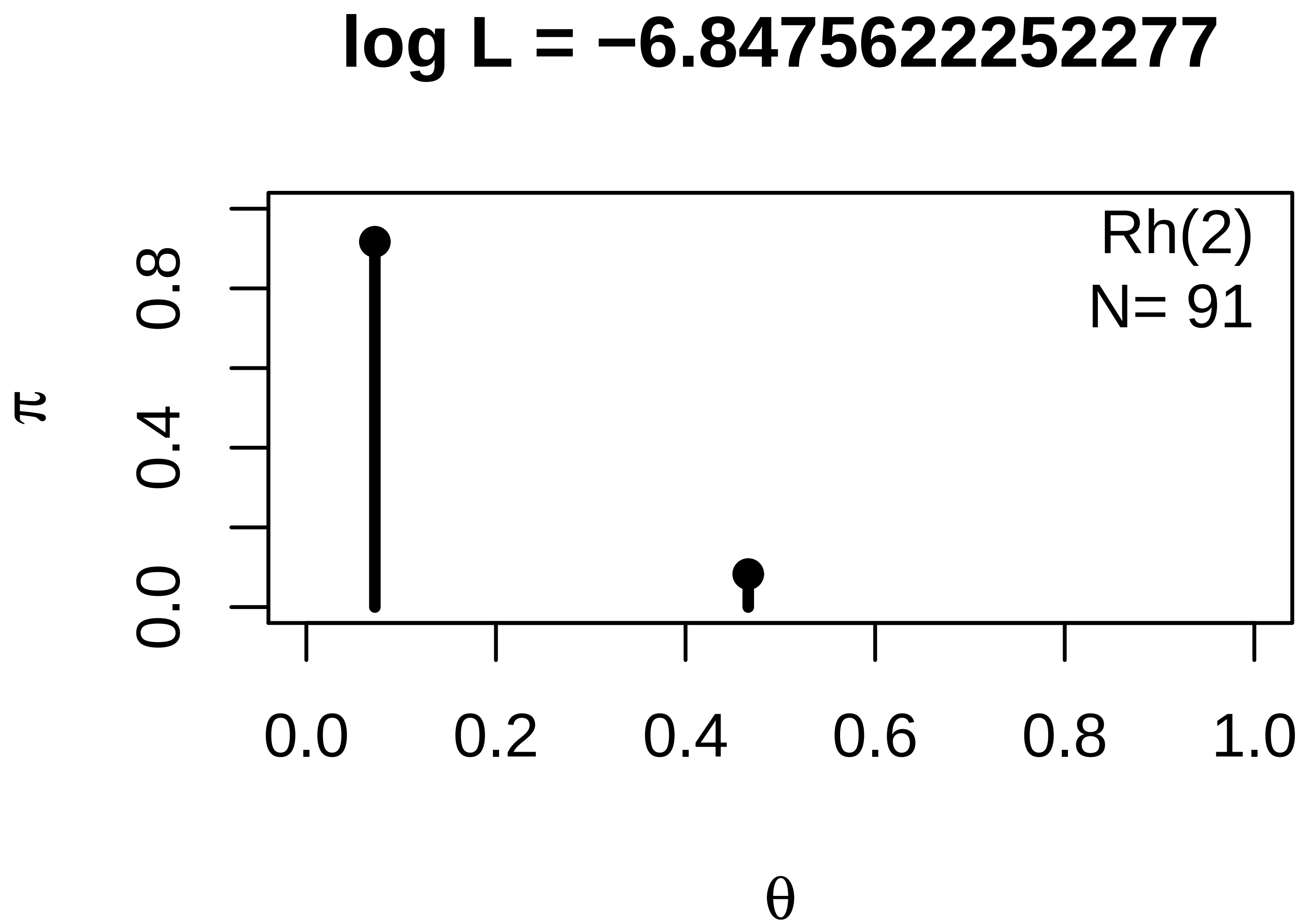
Step 61



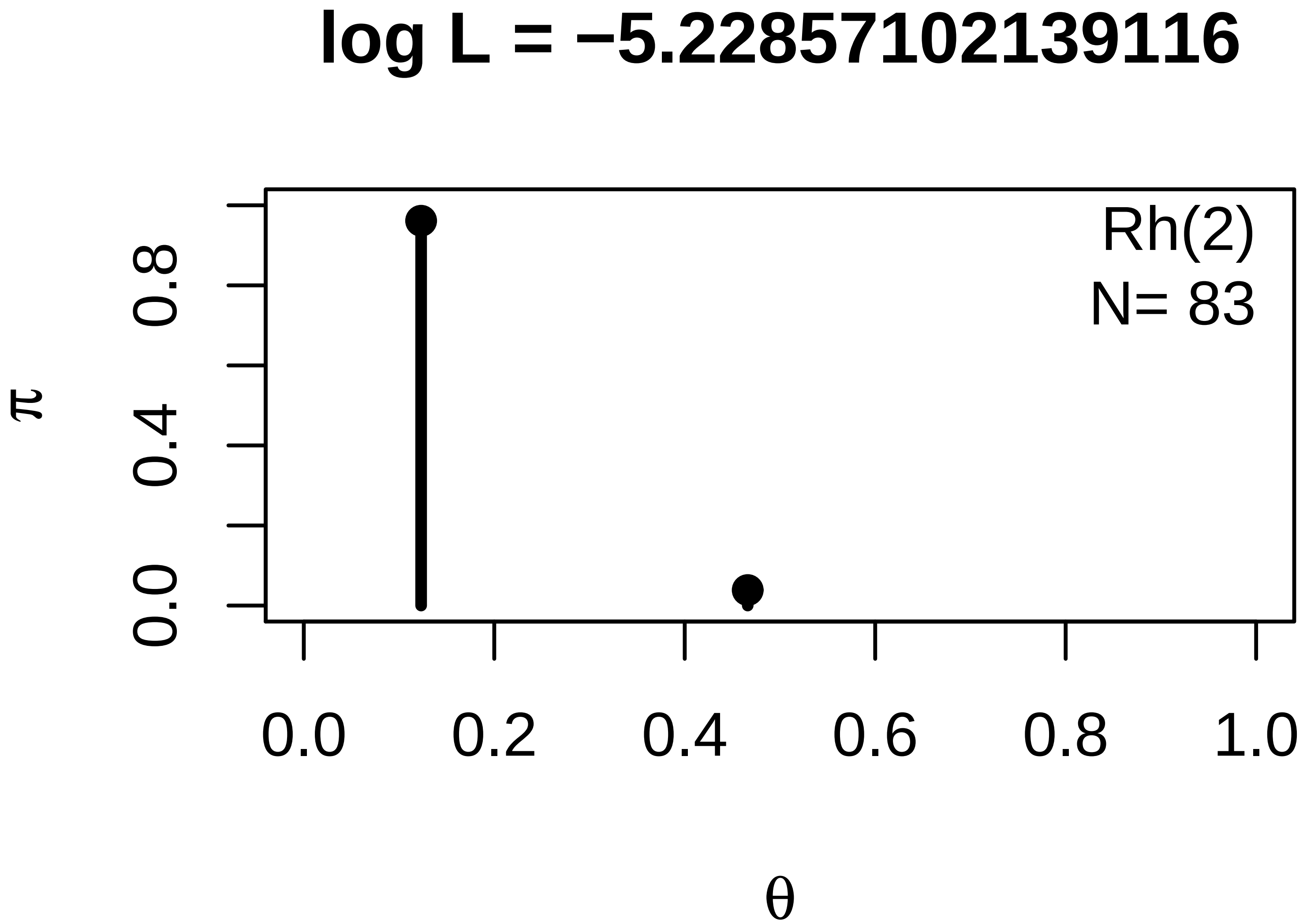
Step 62



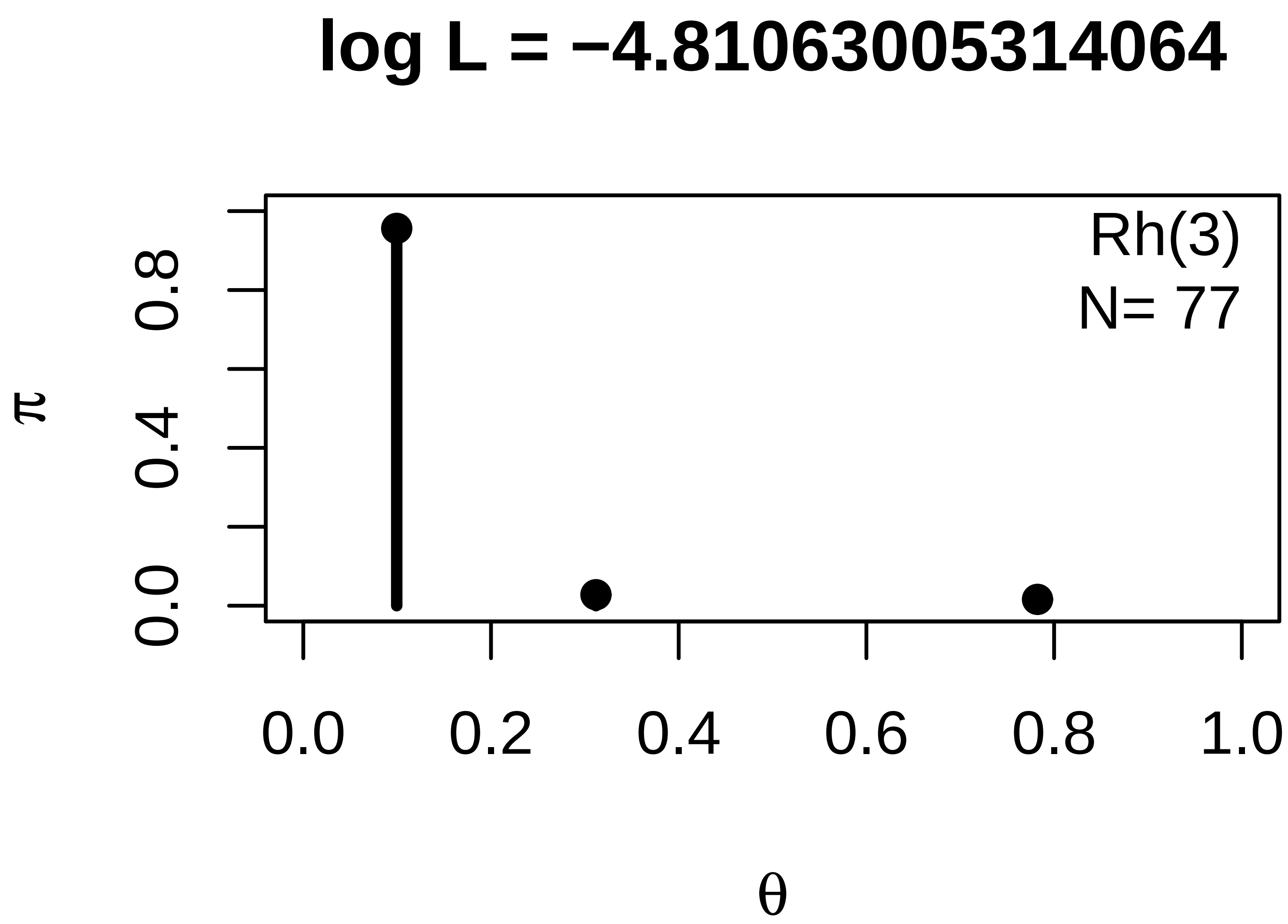
Step 63



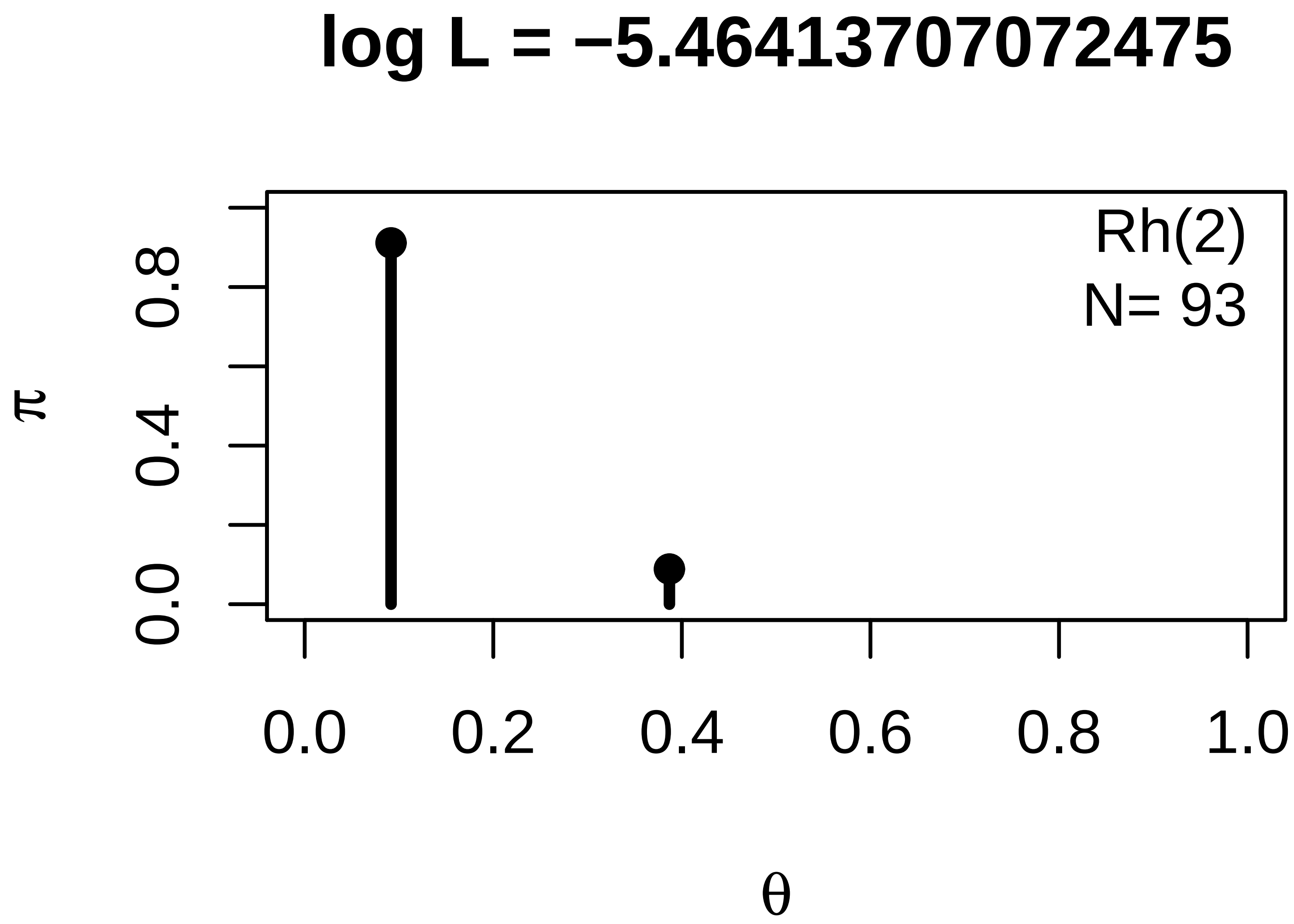
Step 64



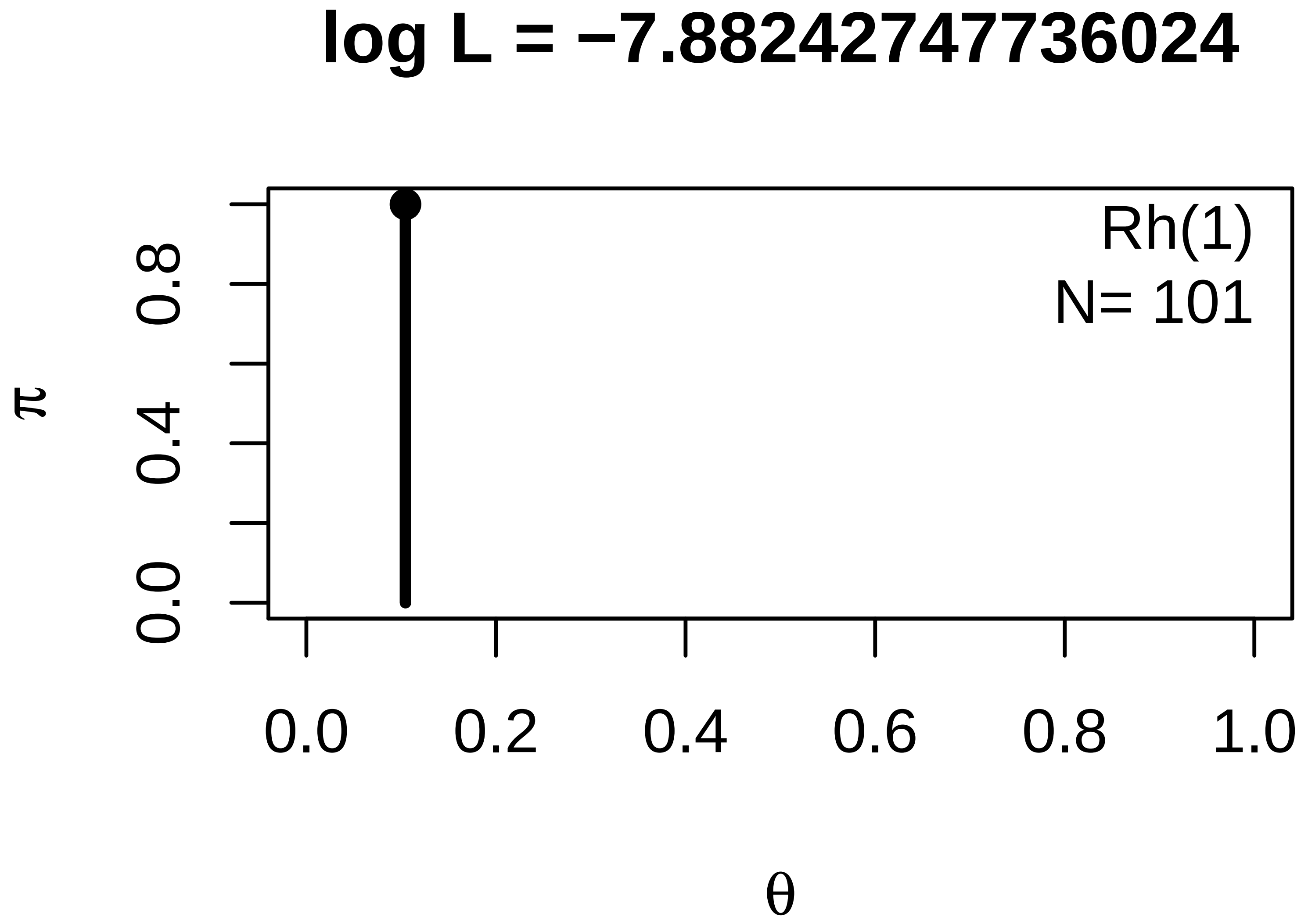
Step 65



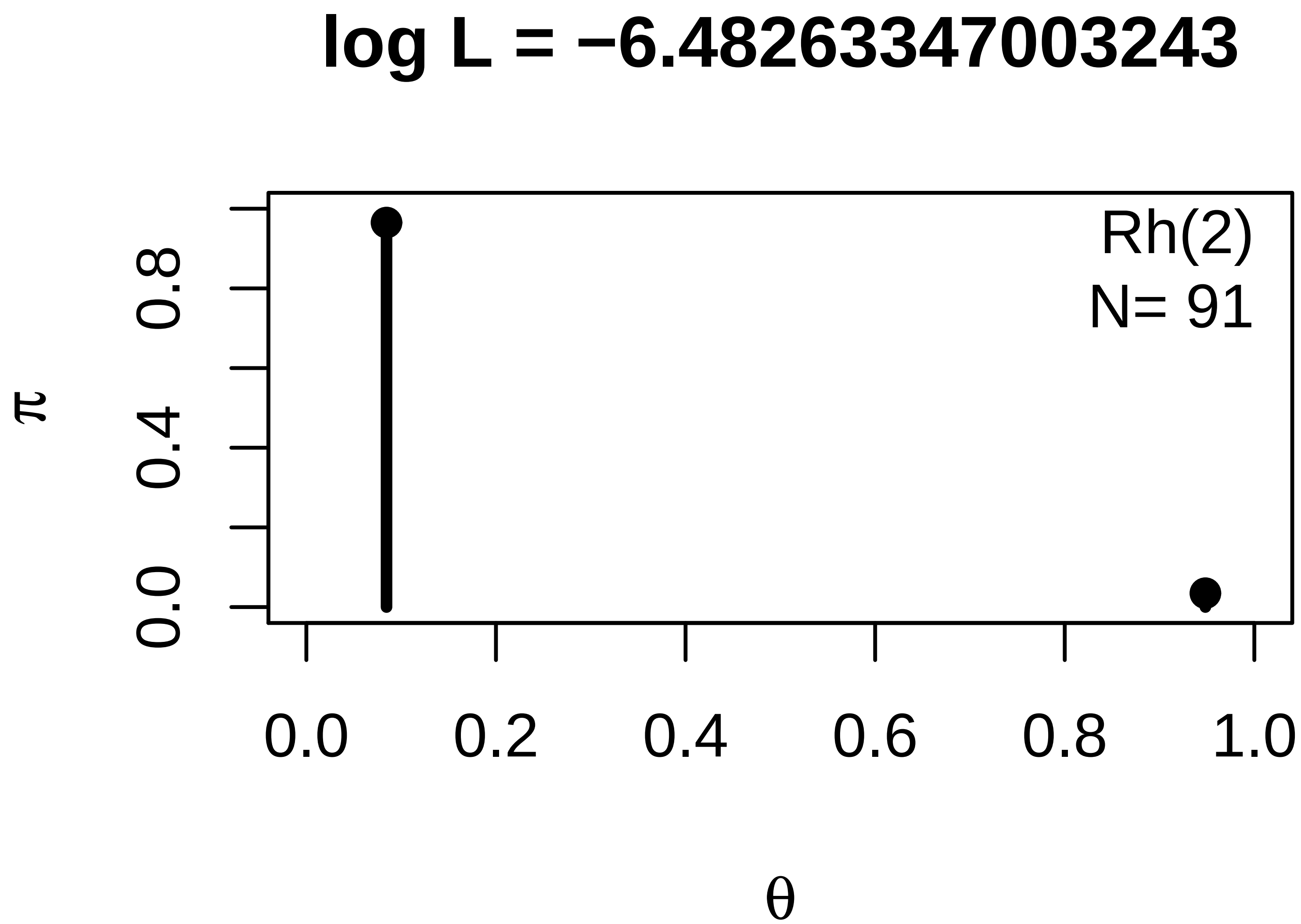
Step 66



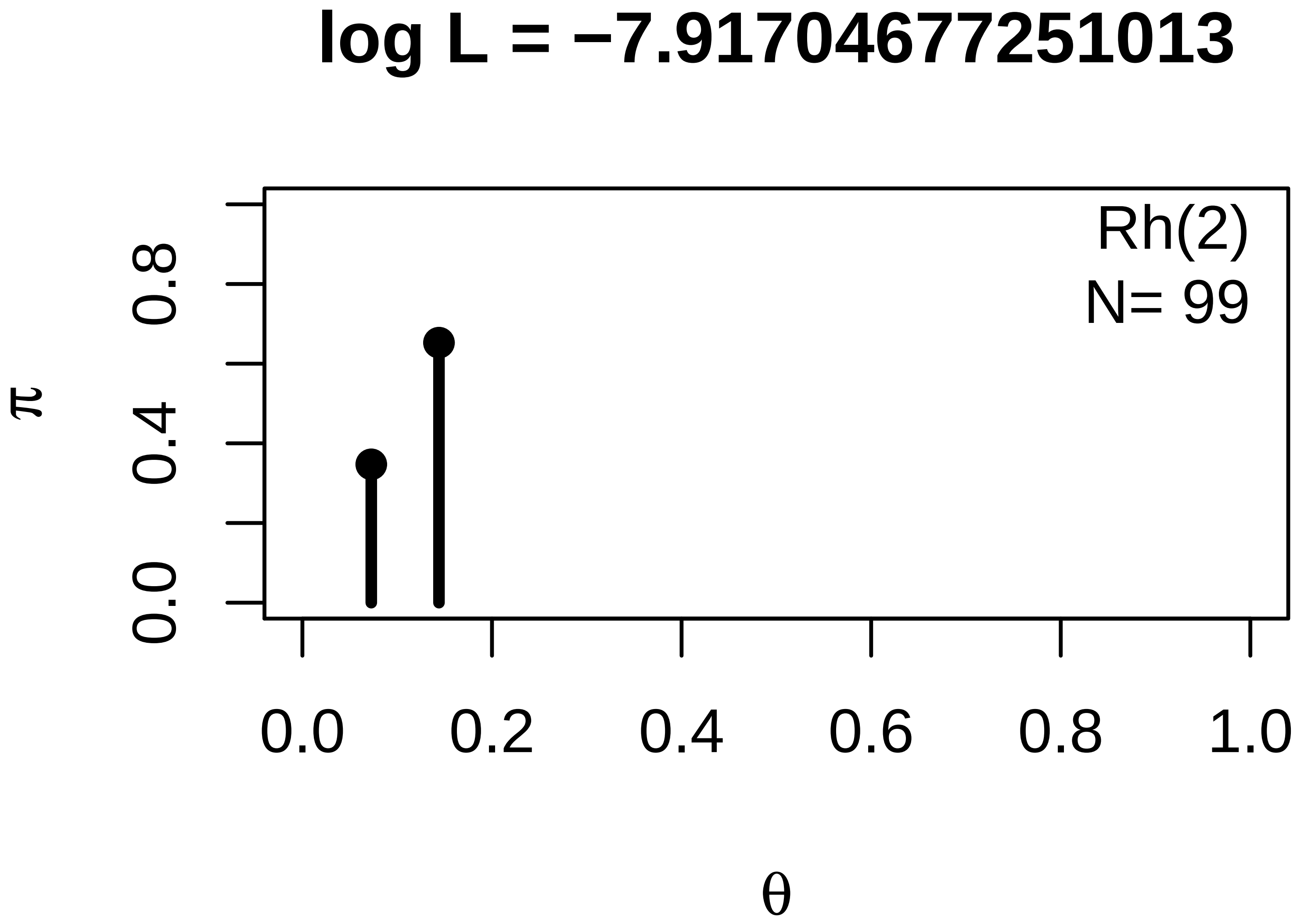
Step 67



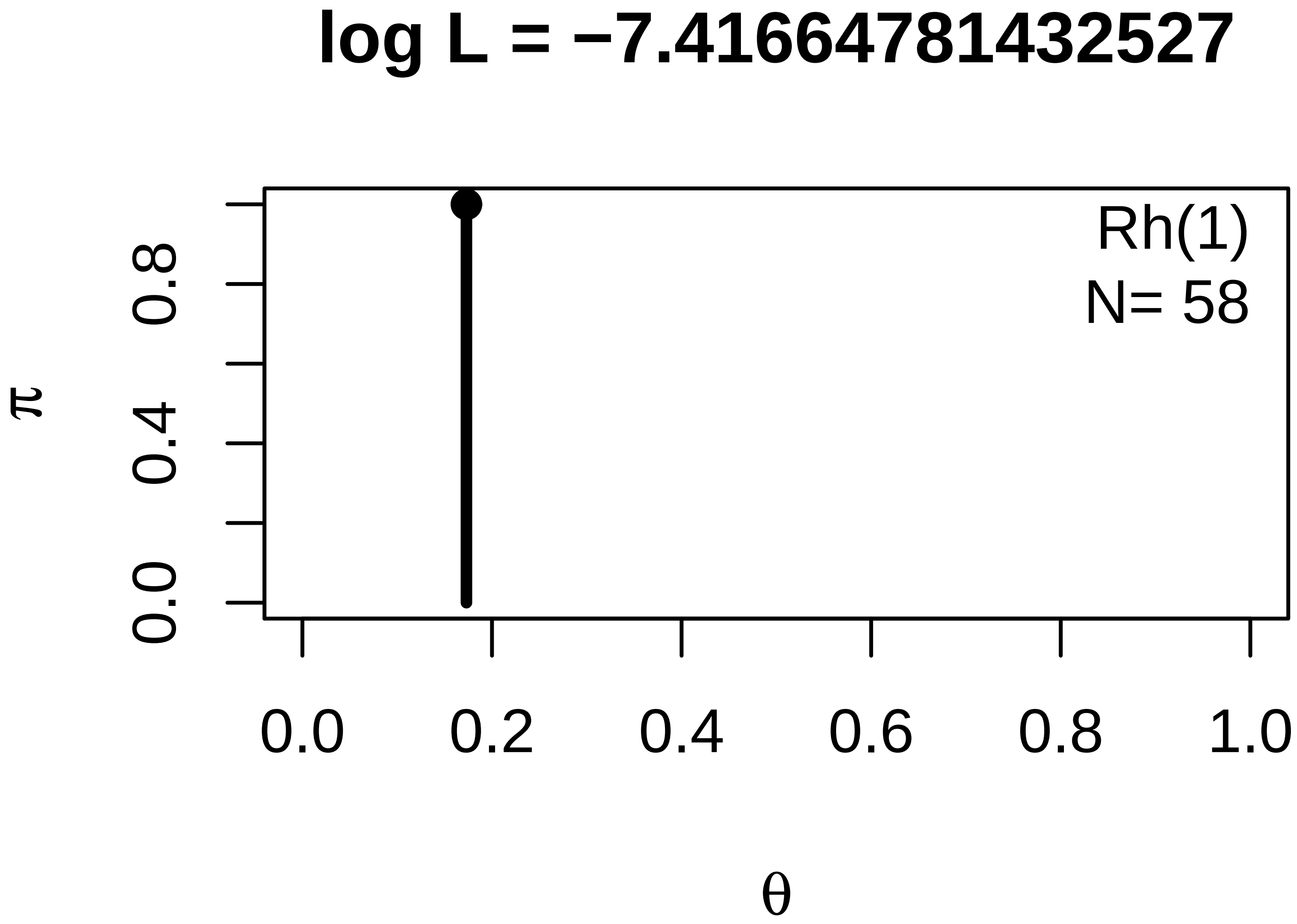
Step 68



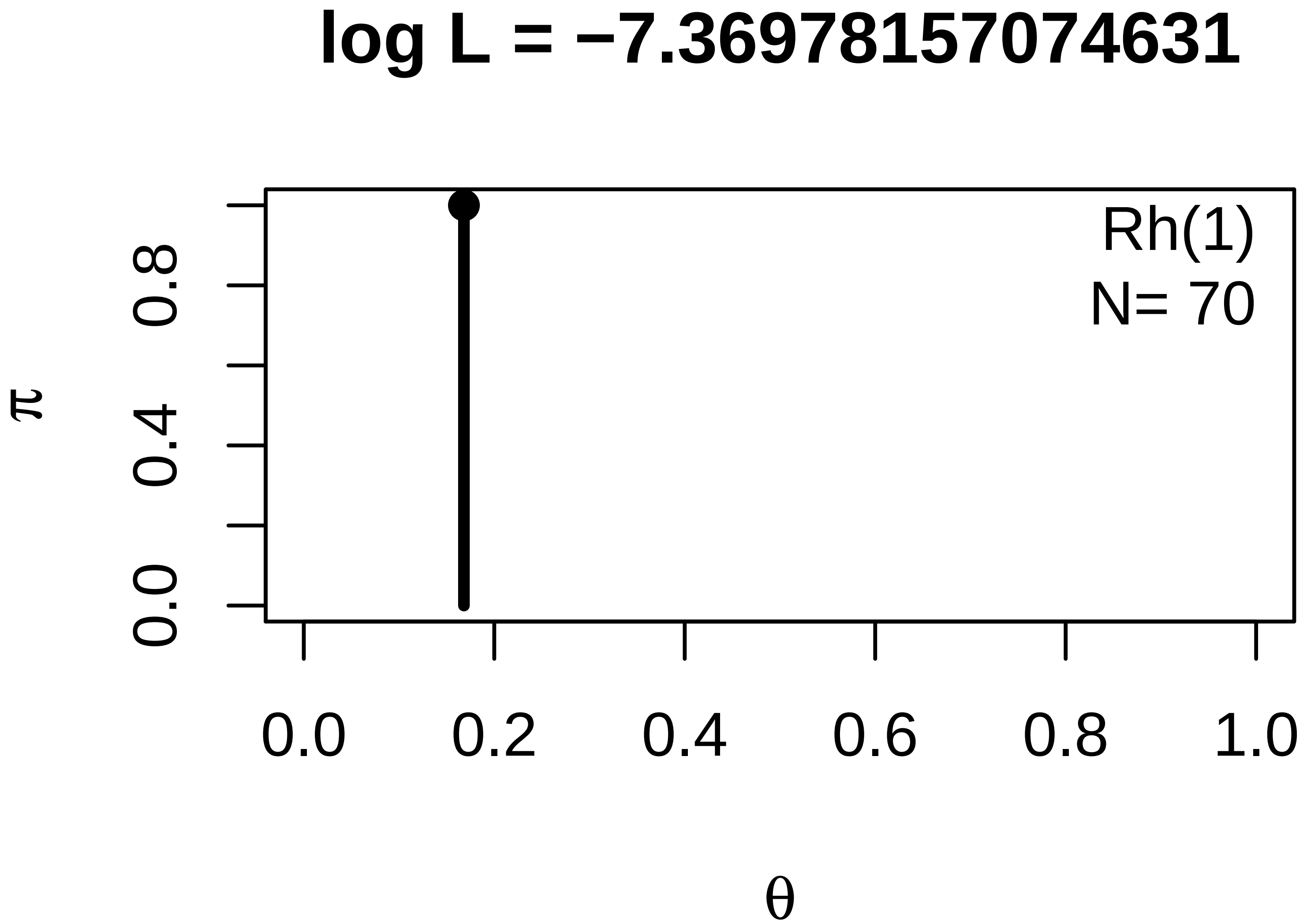
Step 69



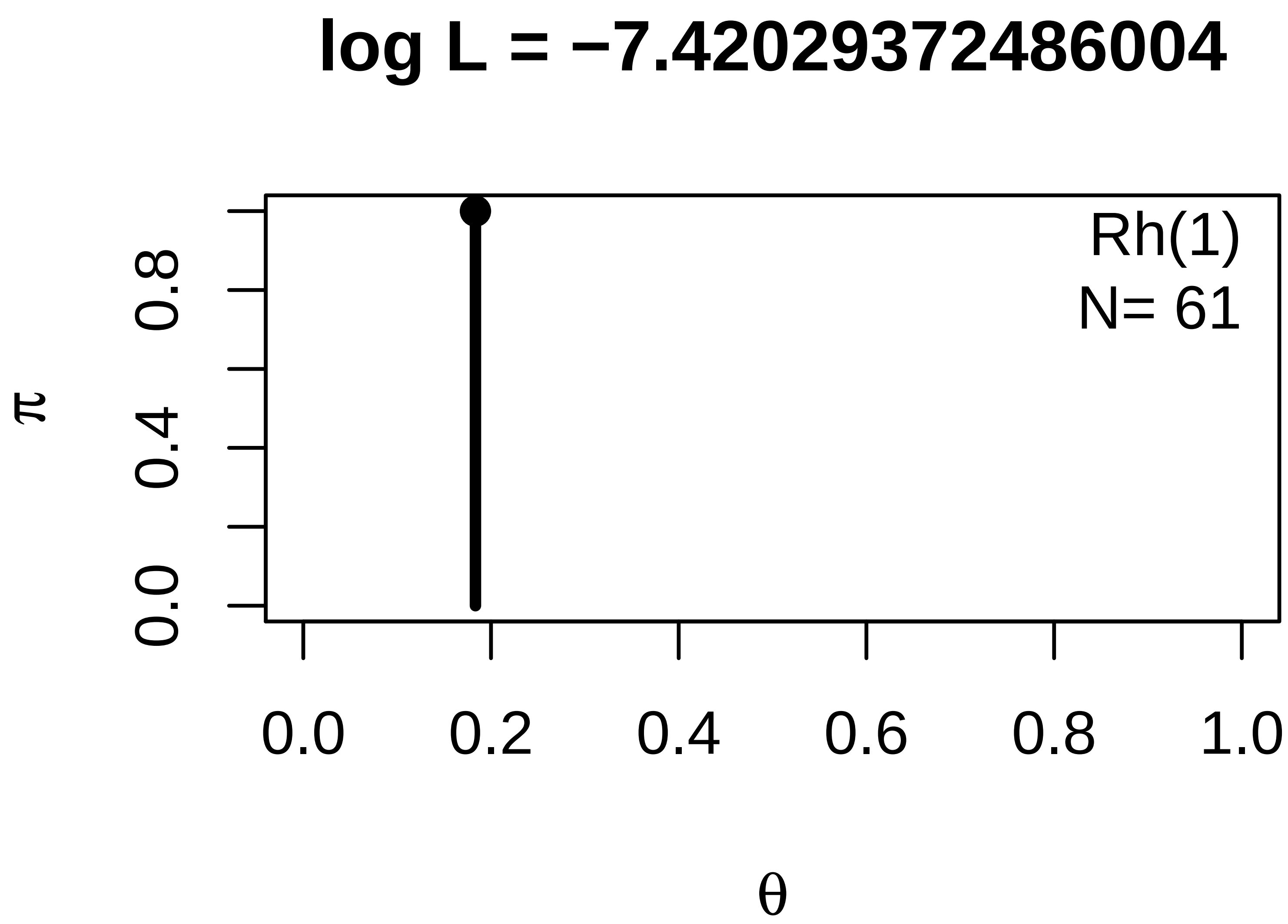
Step 70



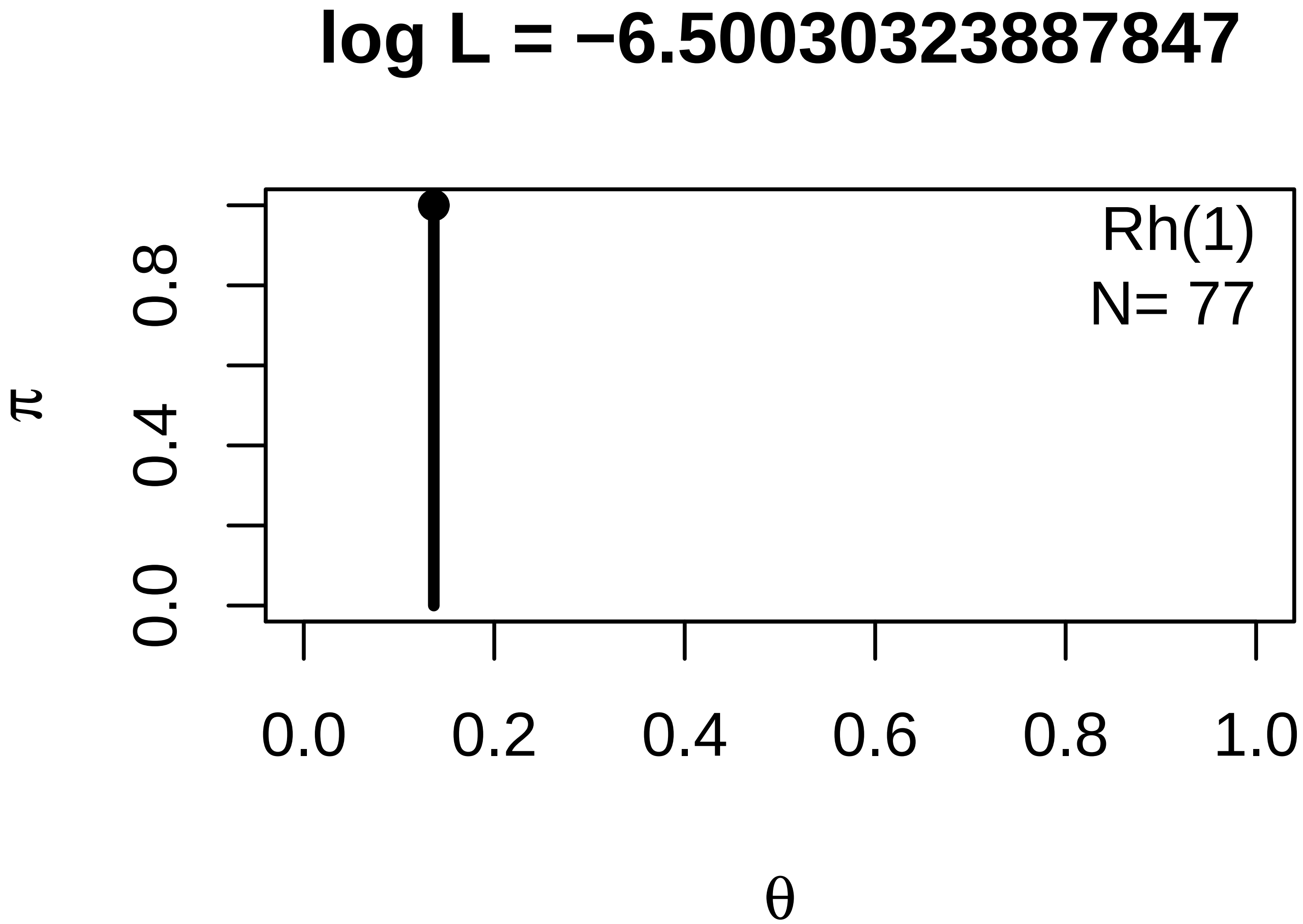
Step 71



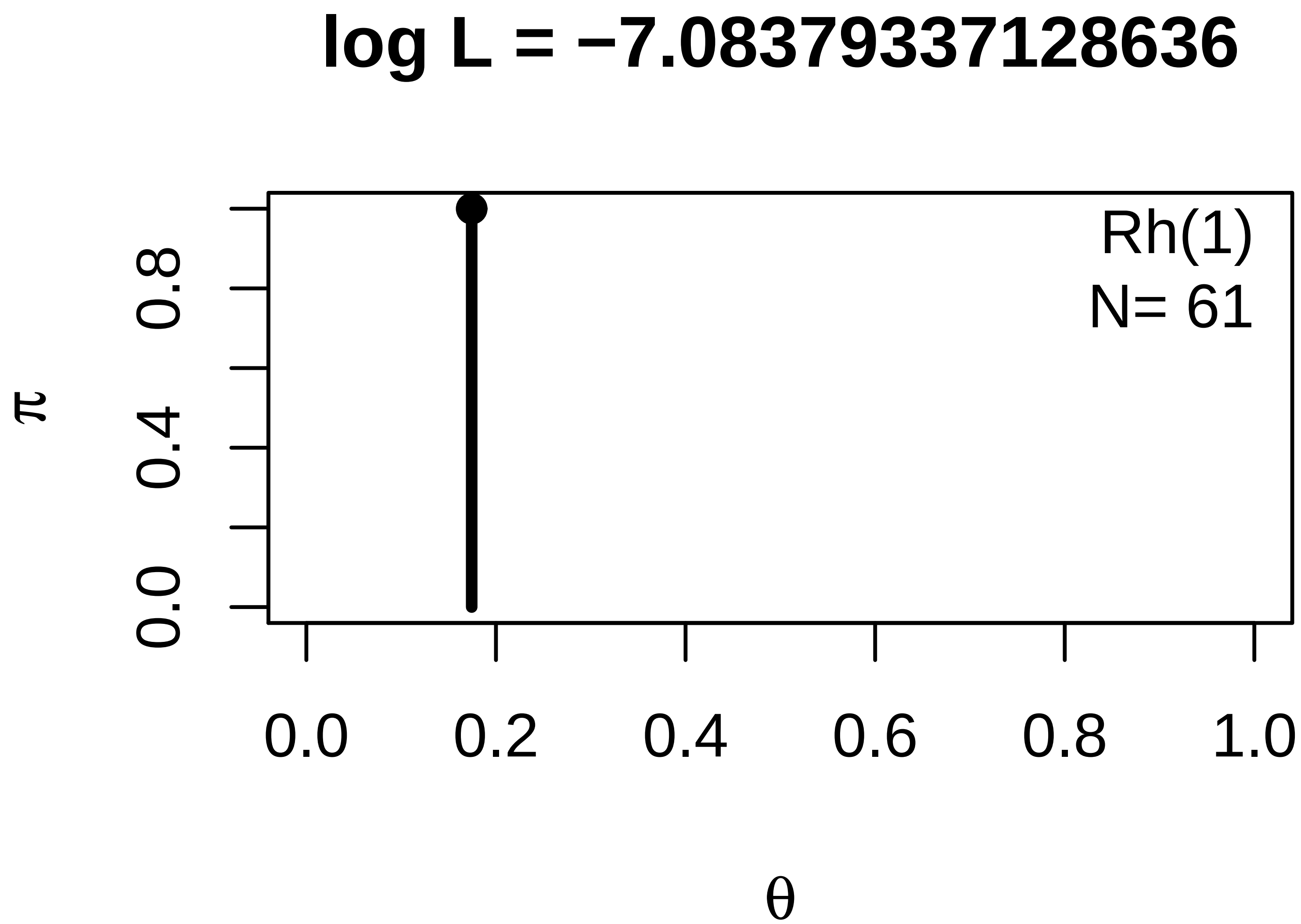
Step 72



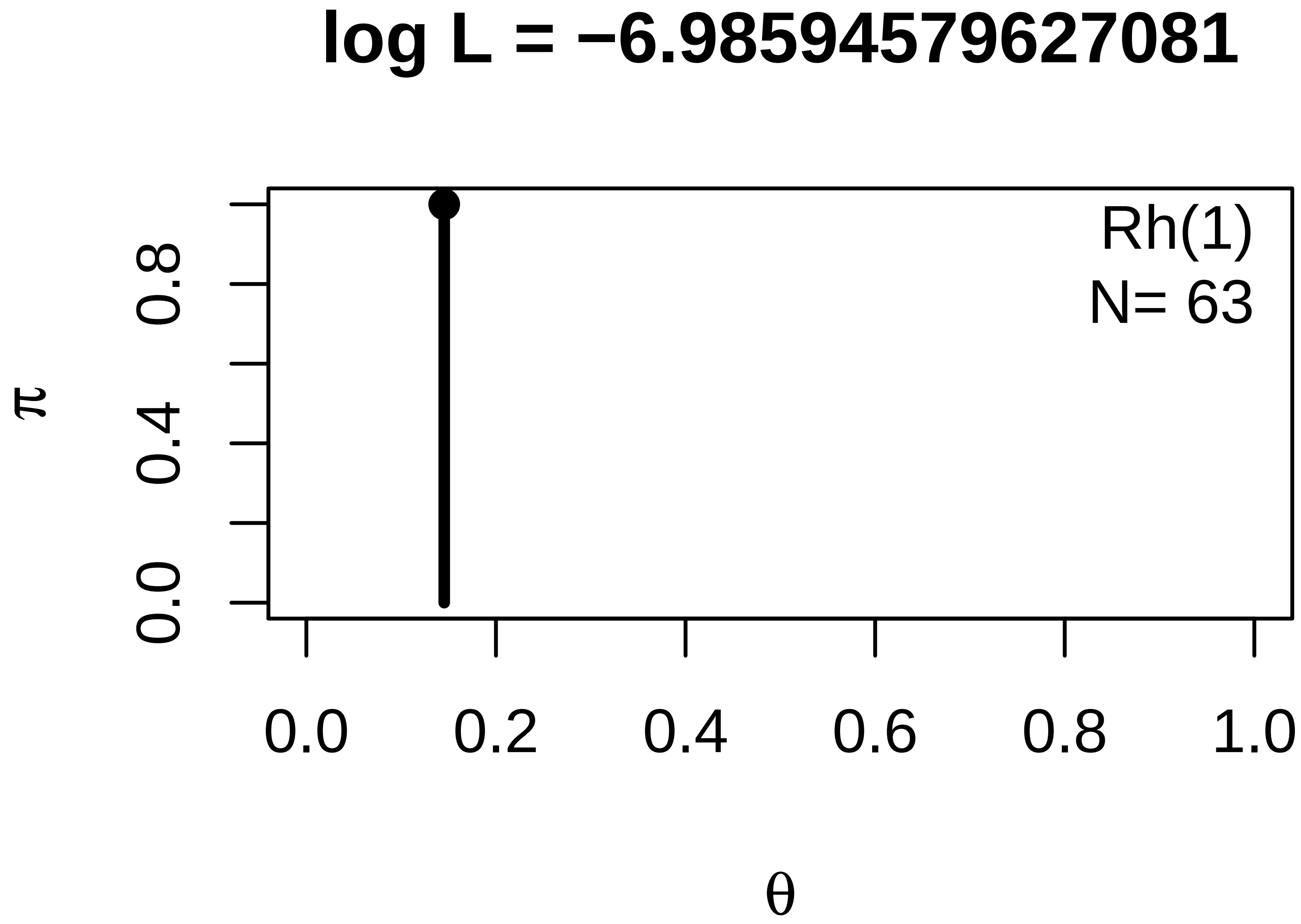
Step 73



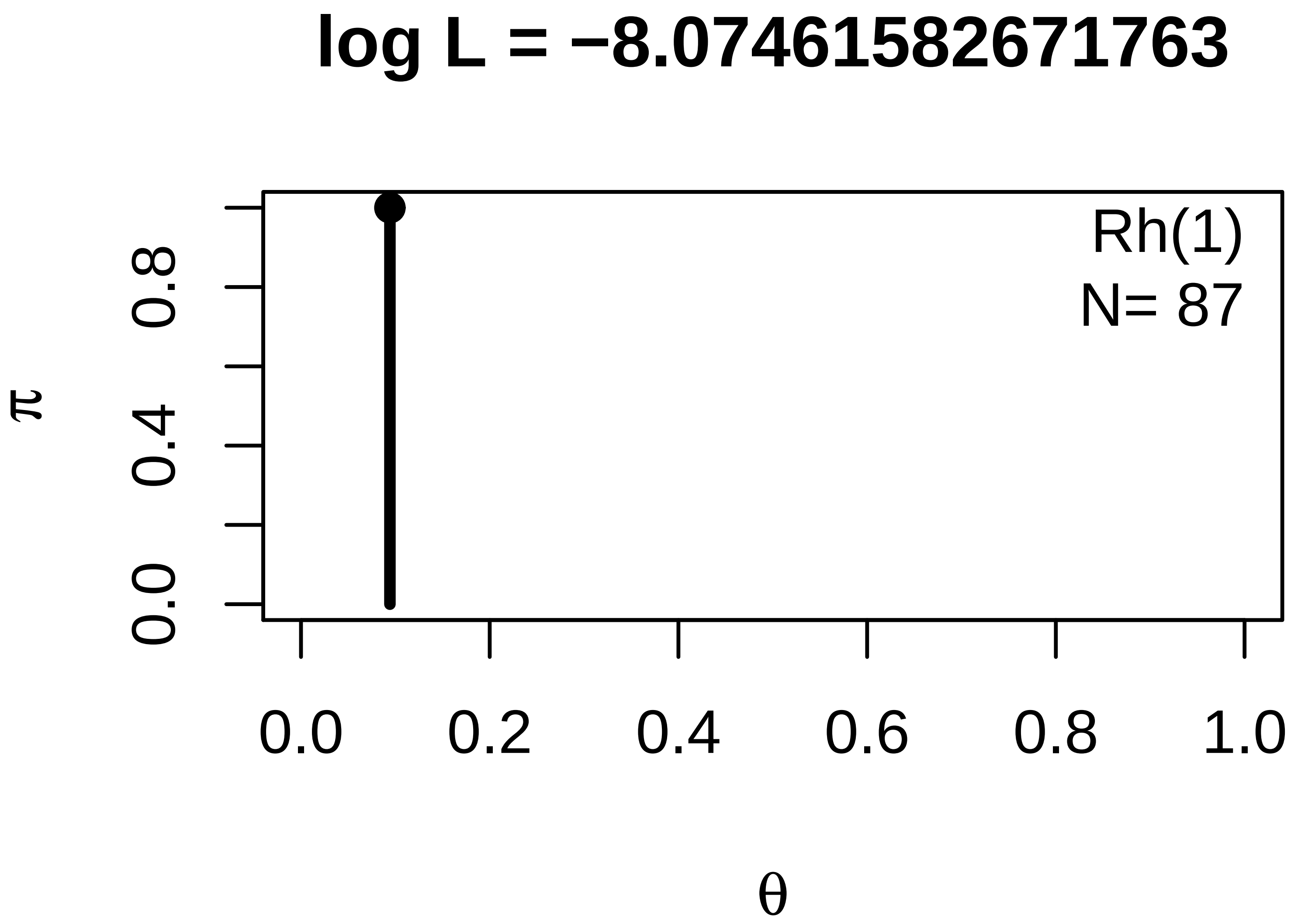
Step 74



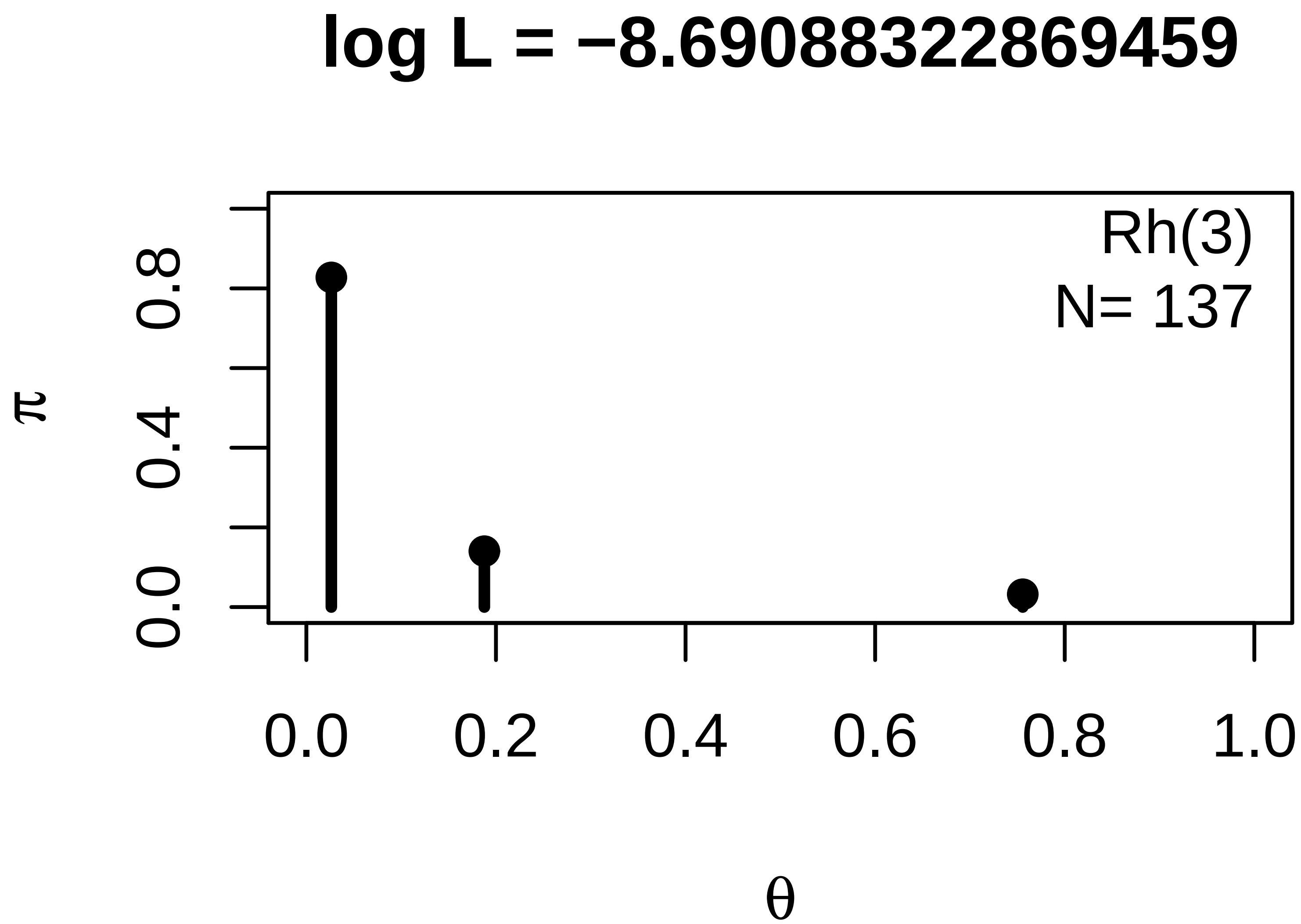
Step 75



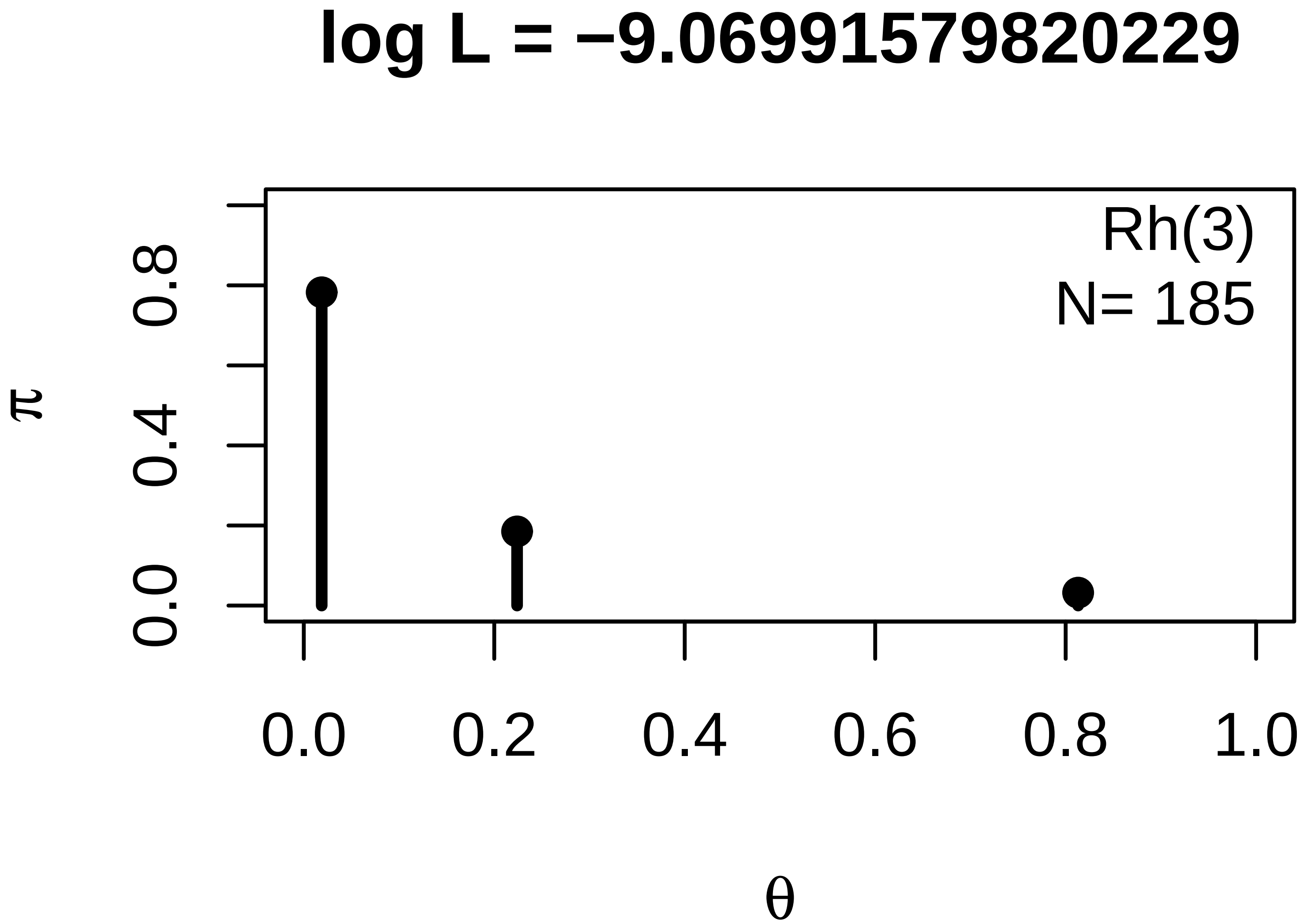
Step 76



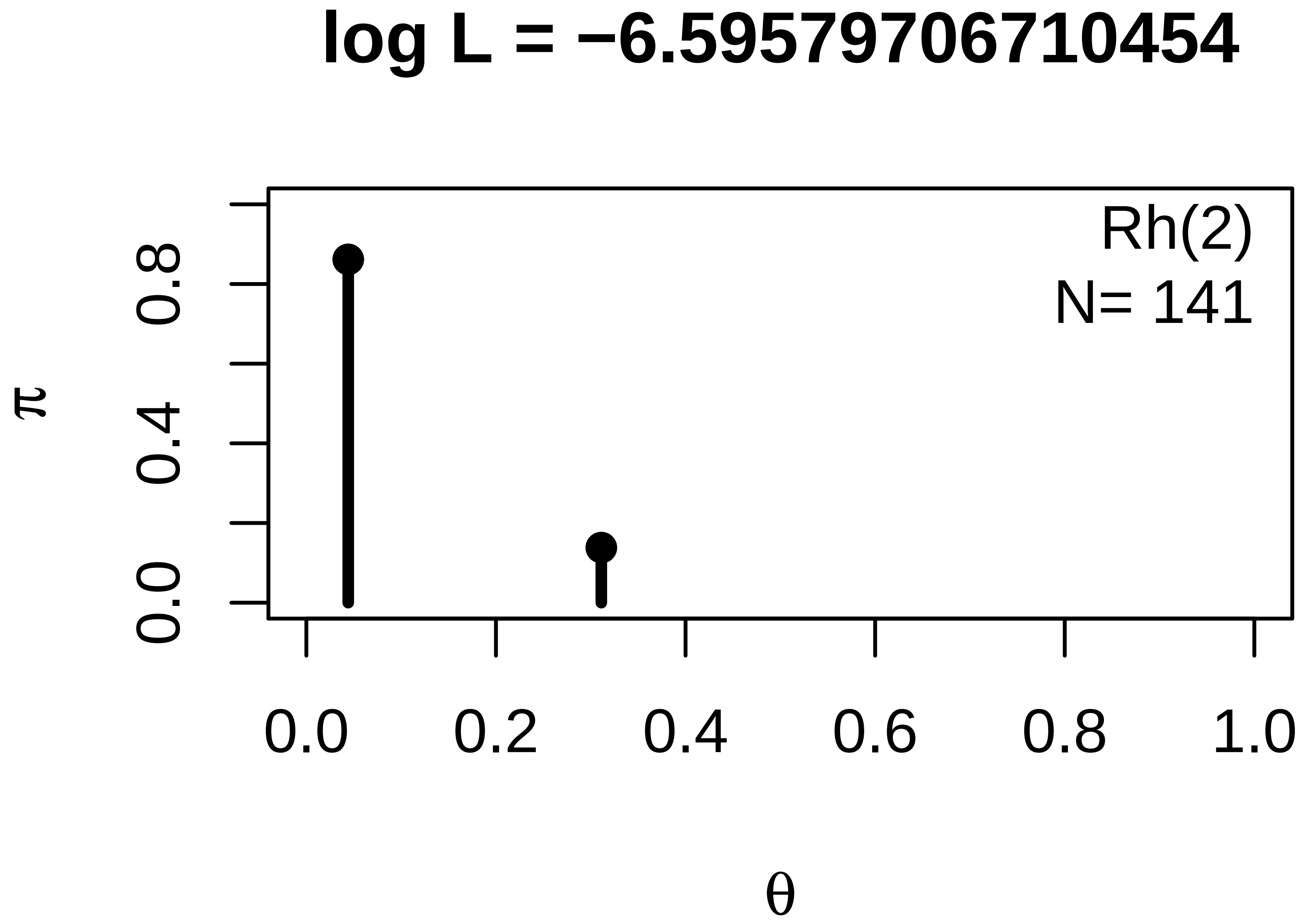
Step 77



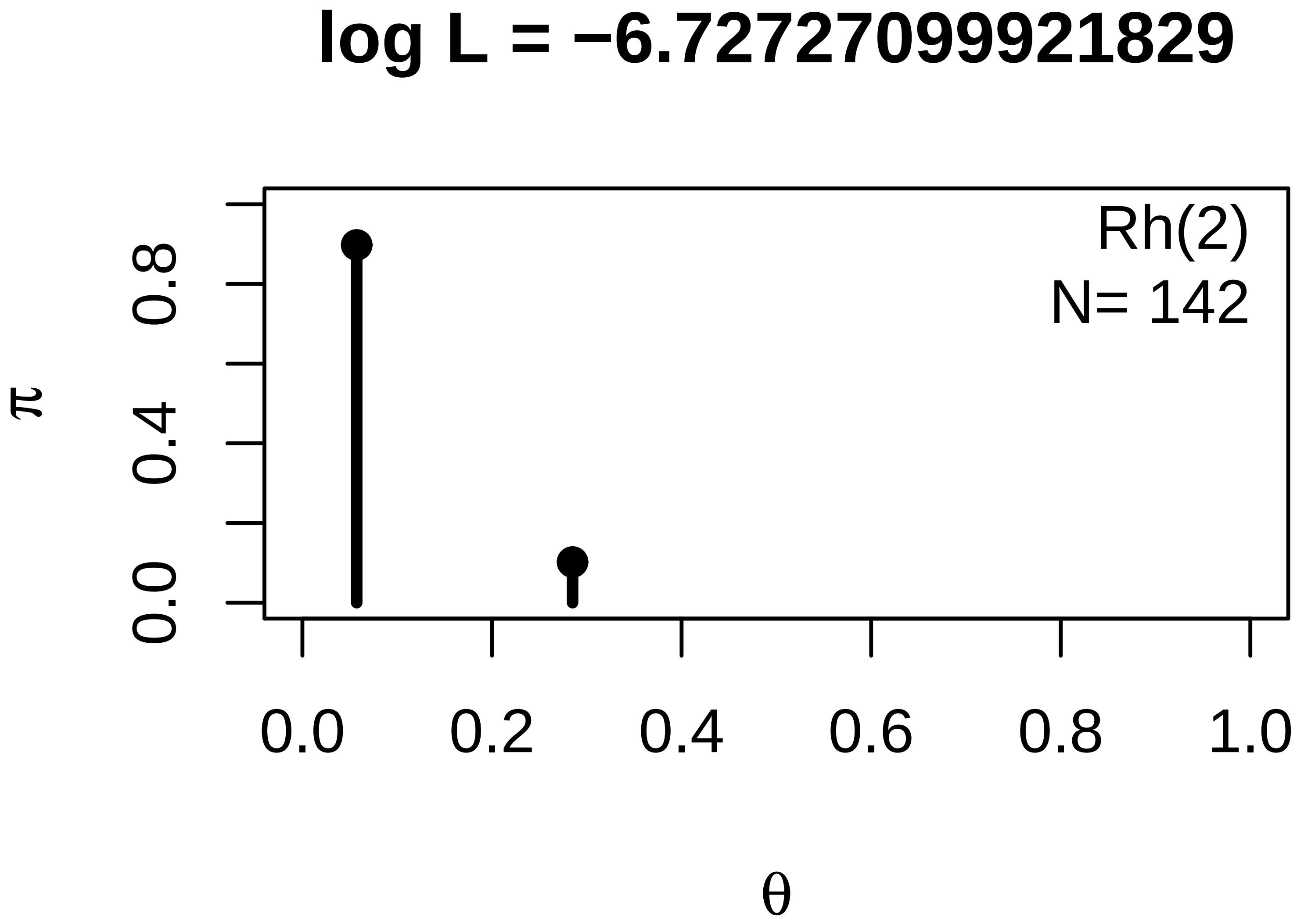
Step 78



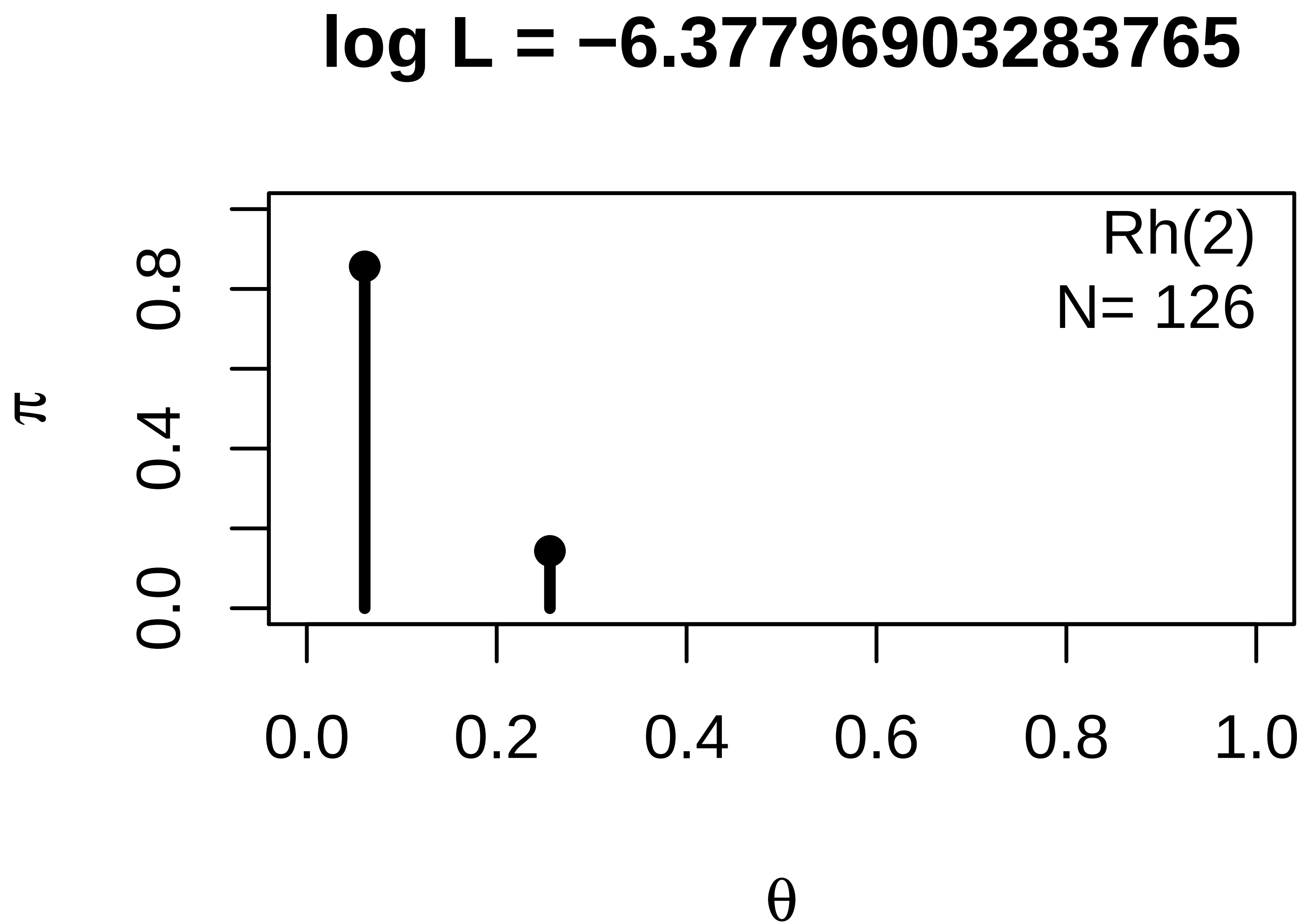
Step 79



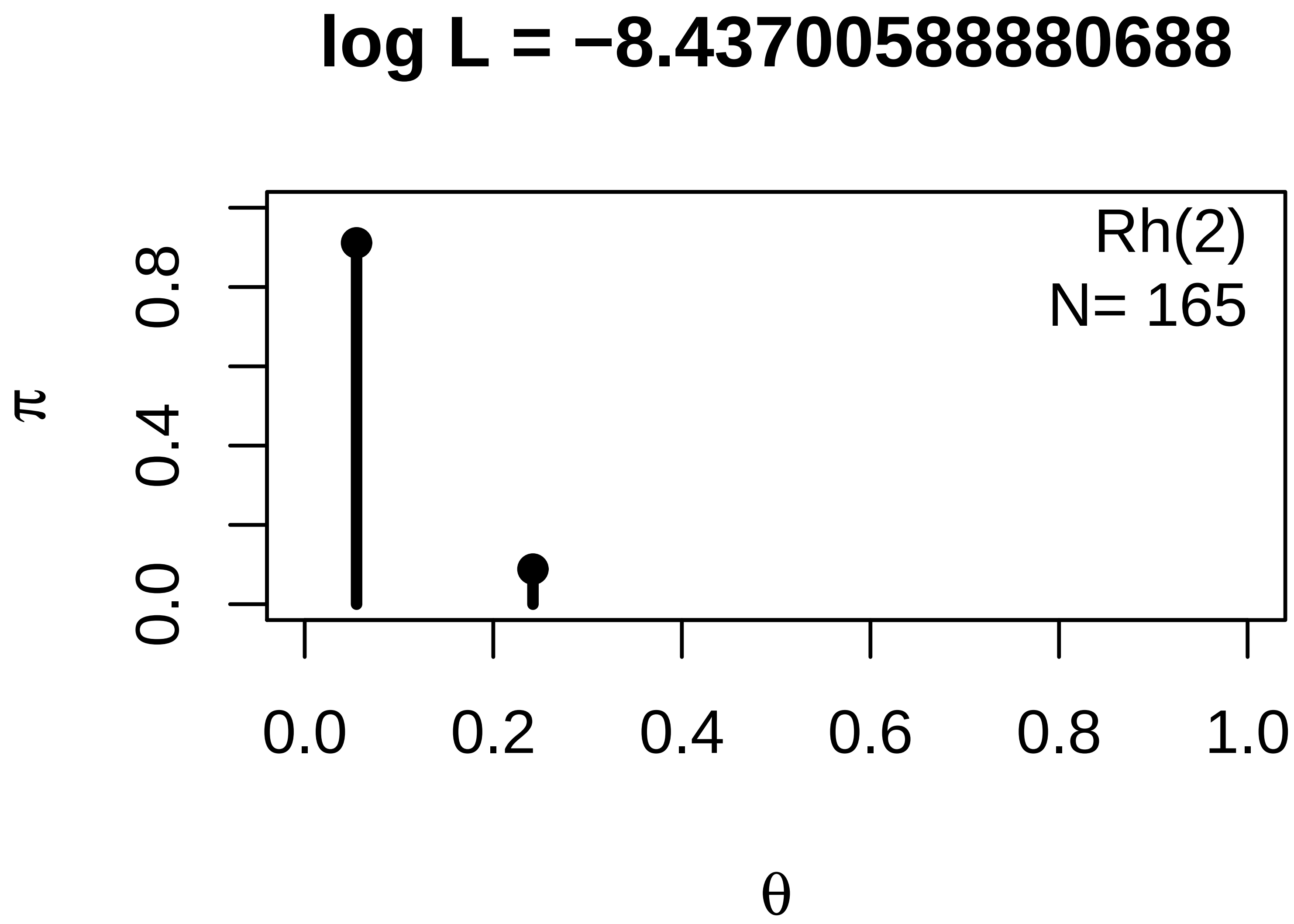
Step 80



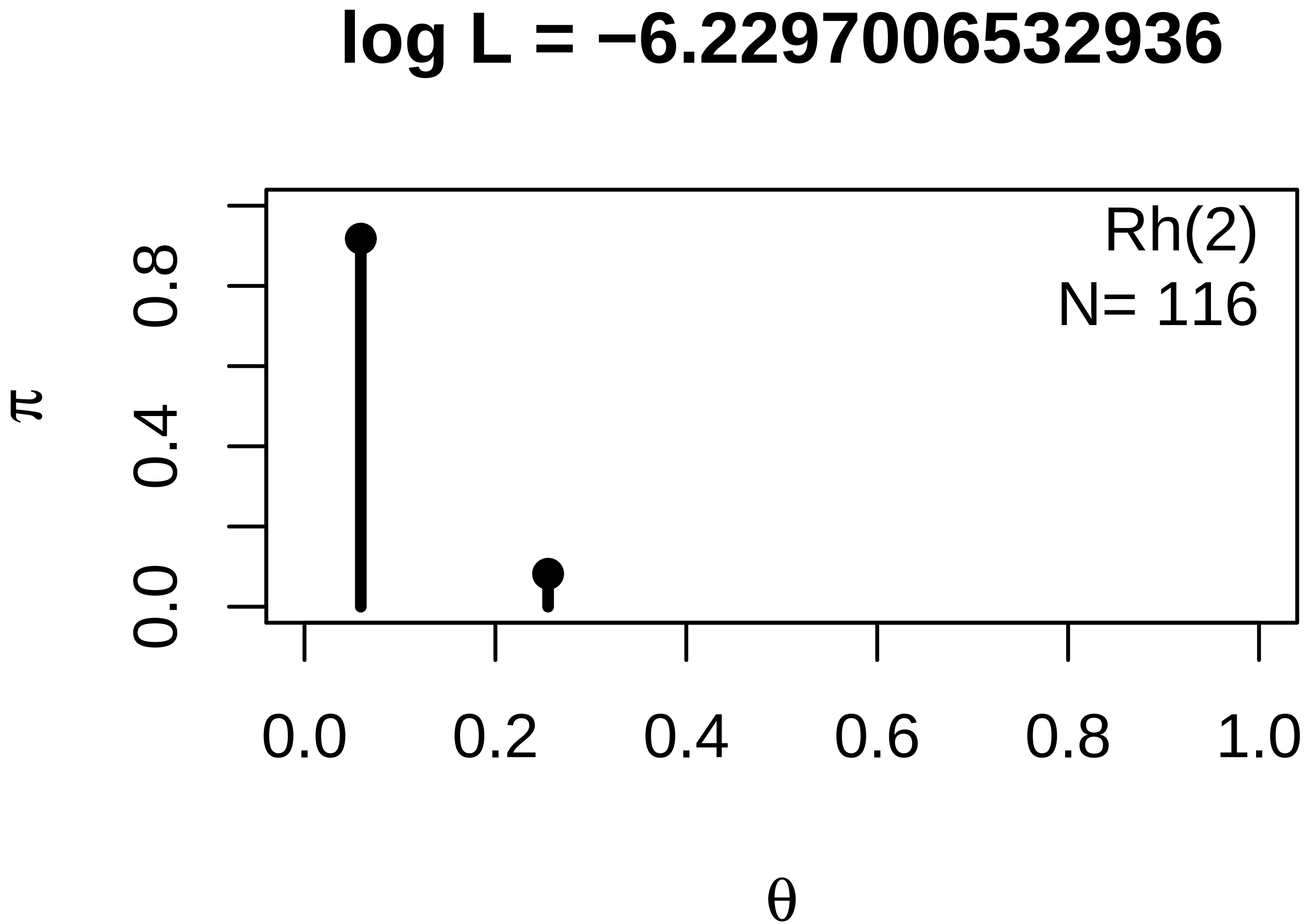
Step 81



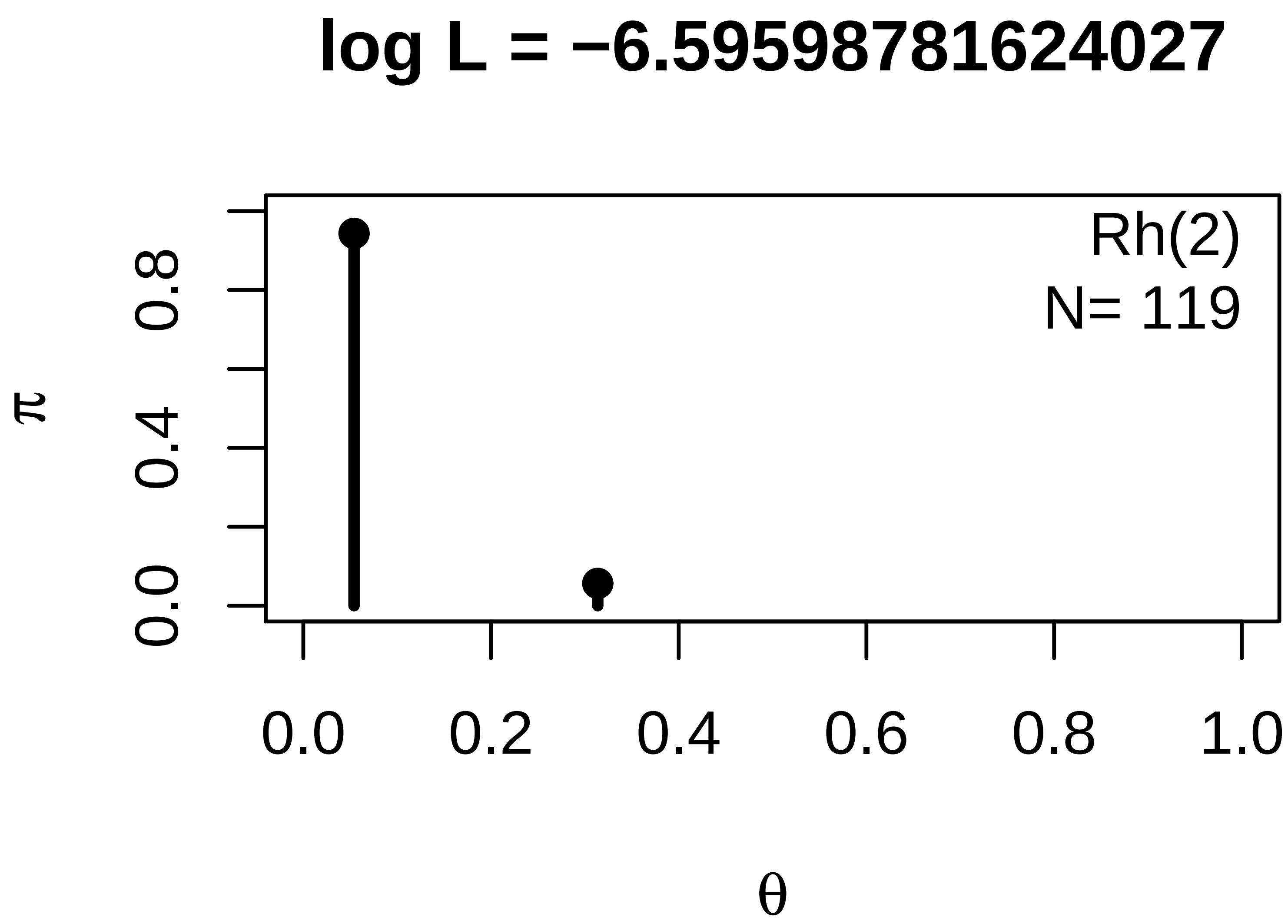
Step 82



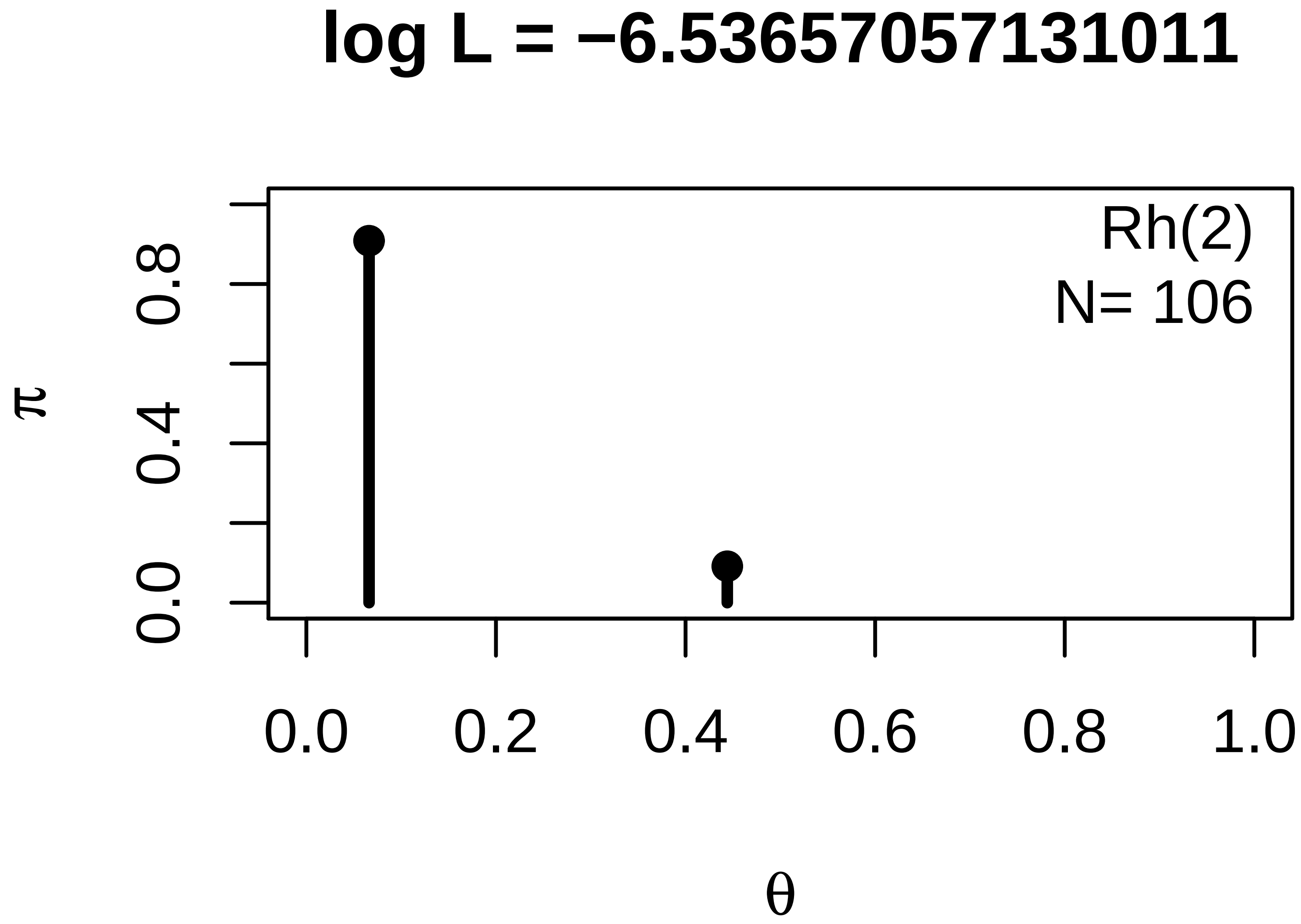
Step 83



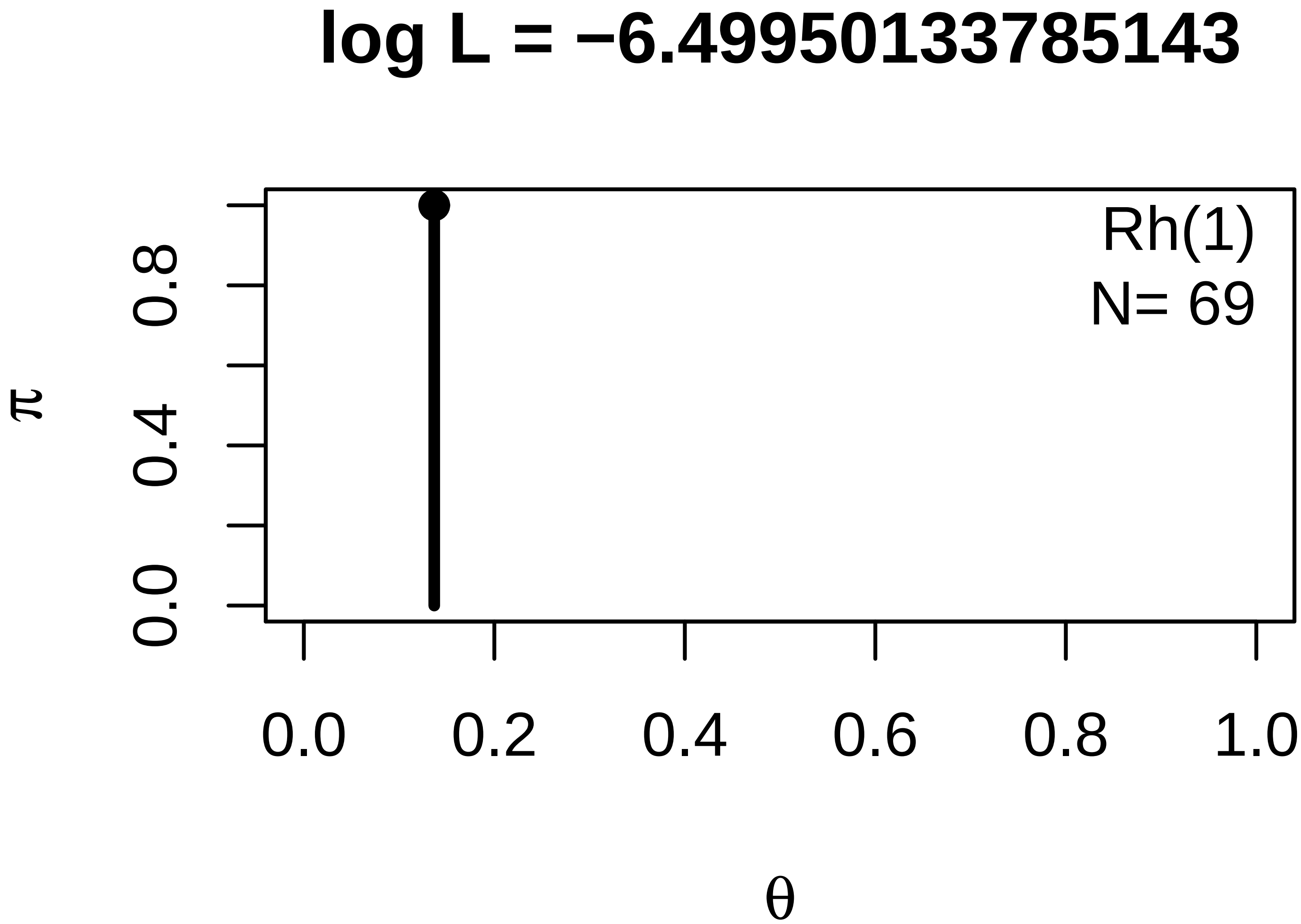
Step 84



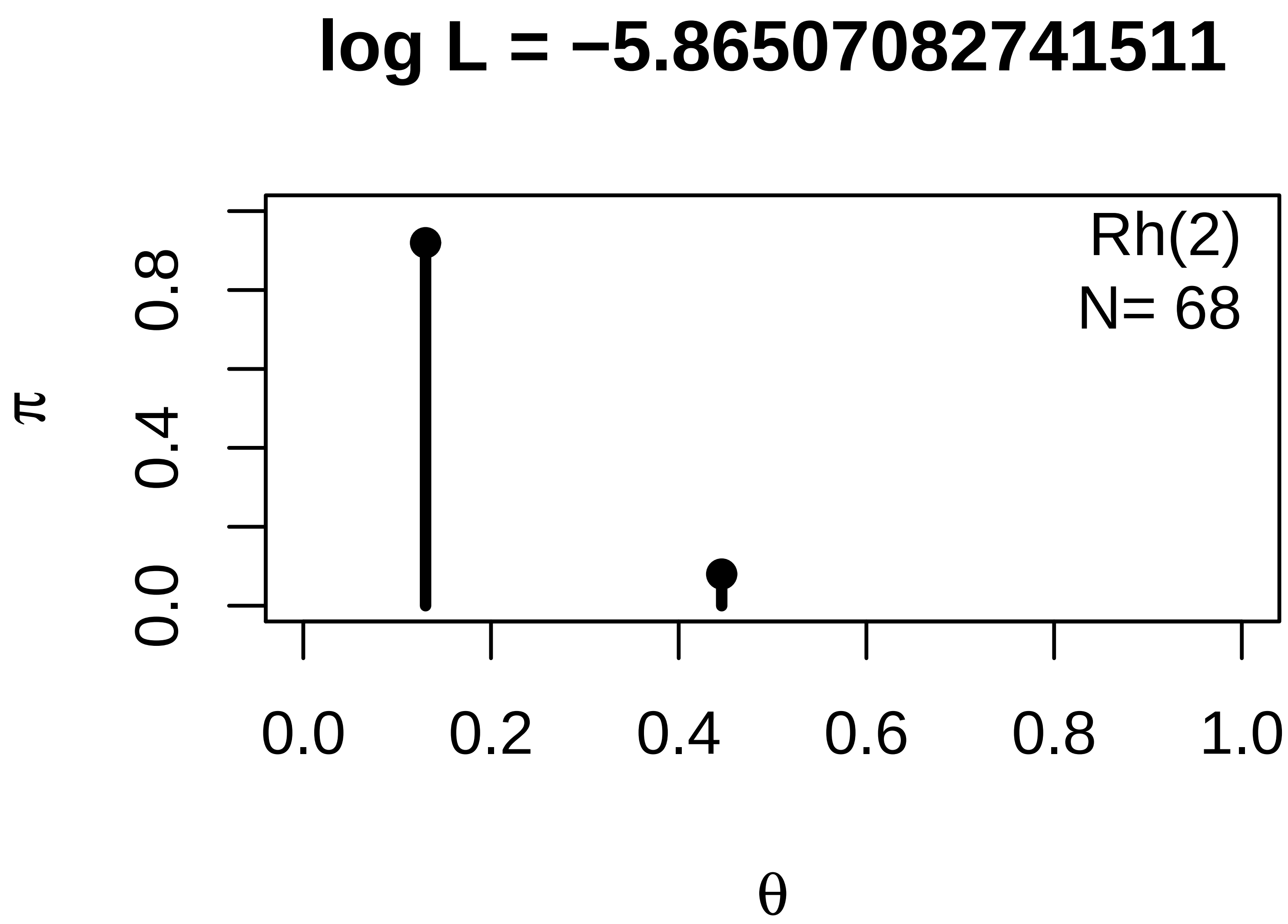
Step 85



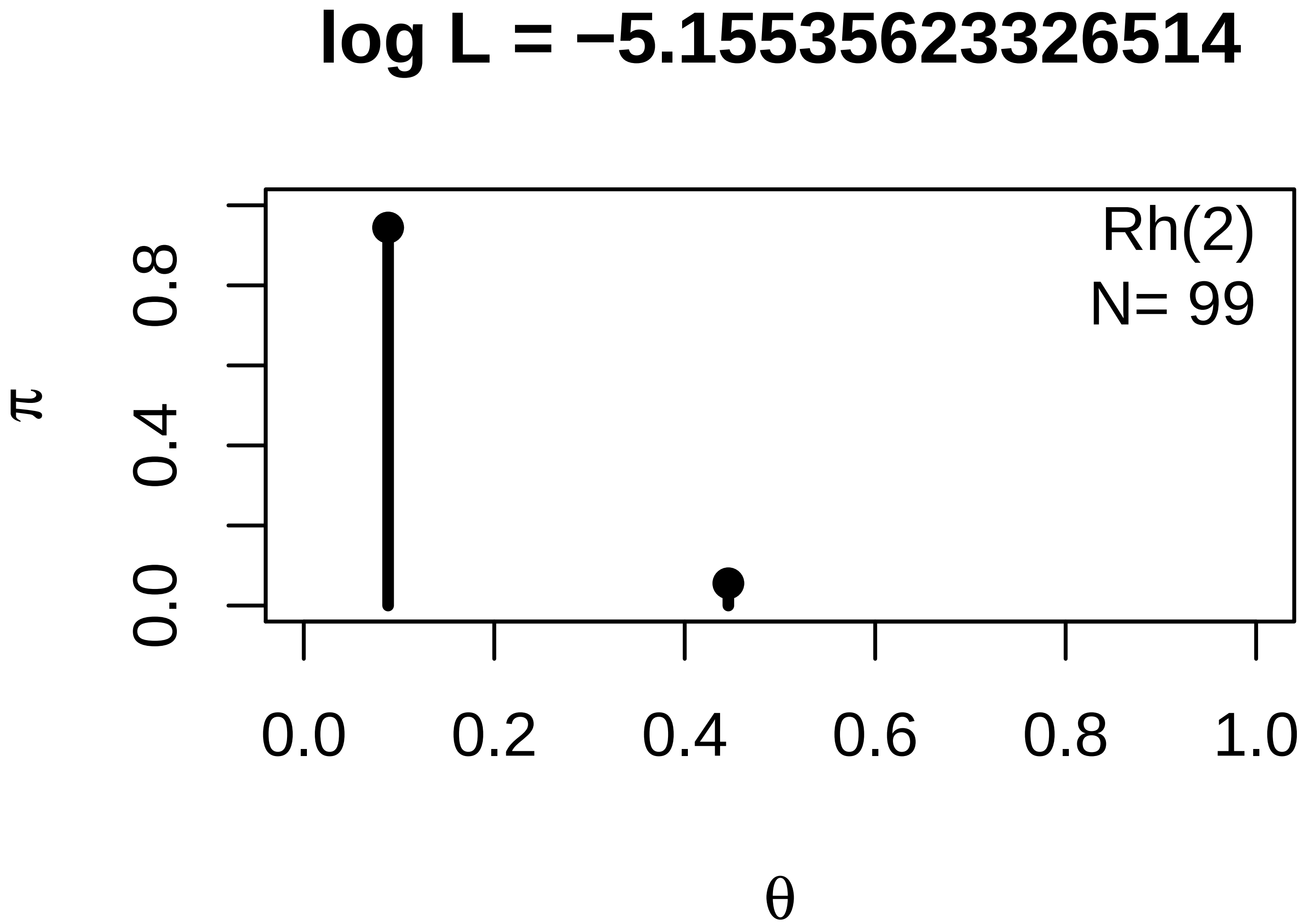
Step 86



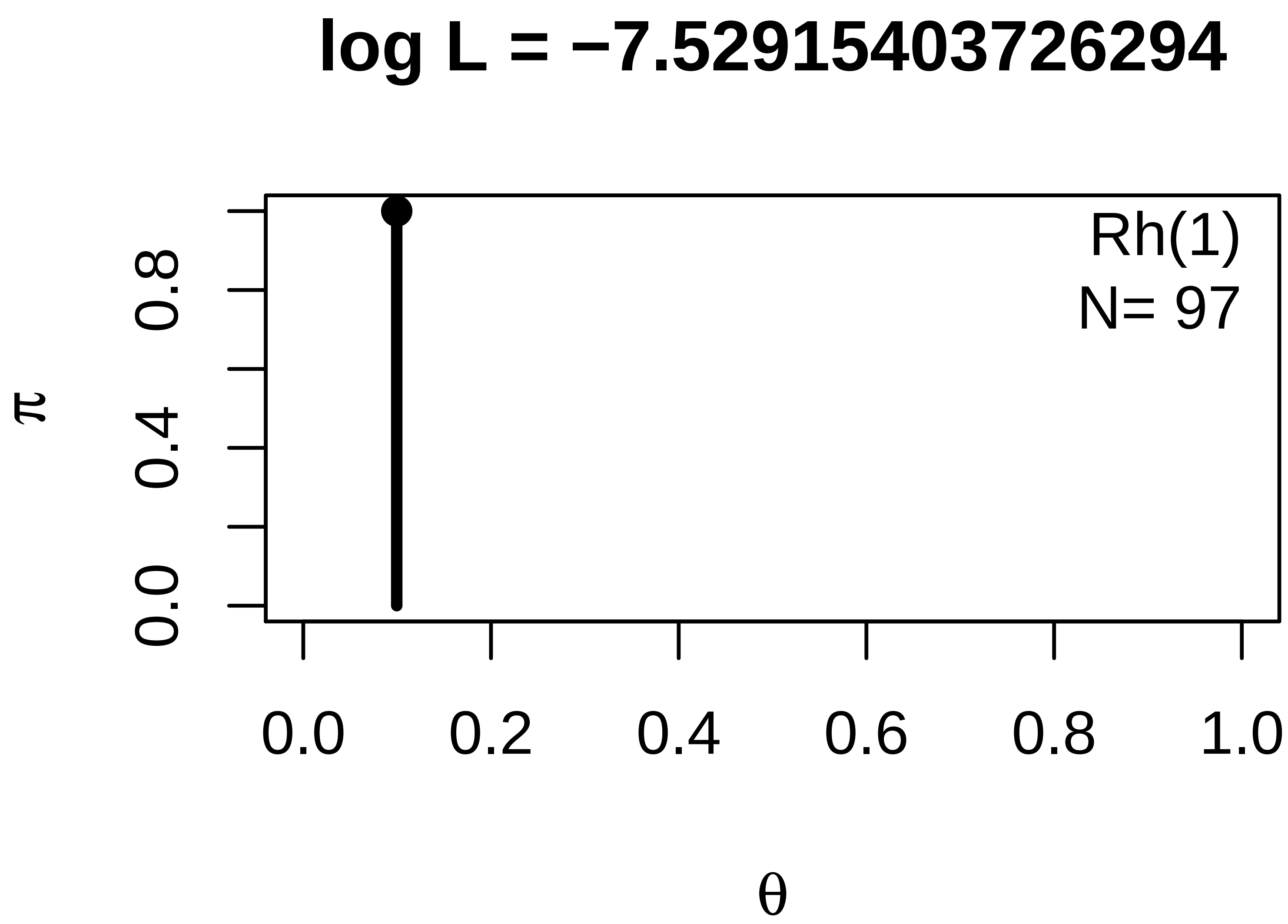
Step 87



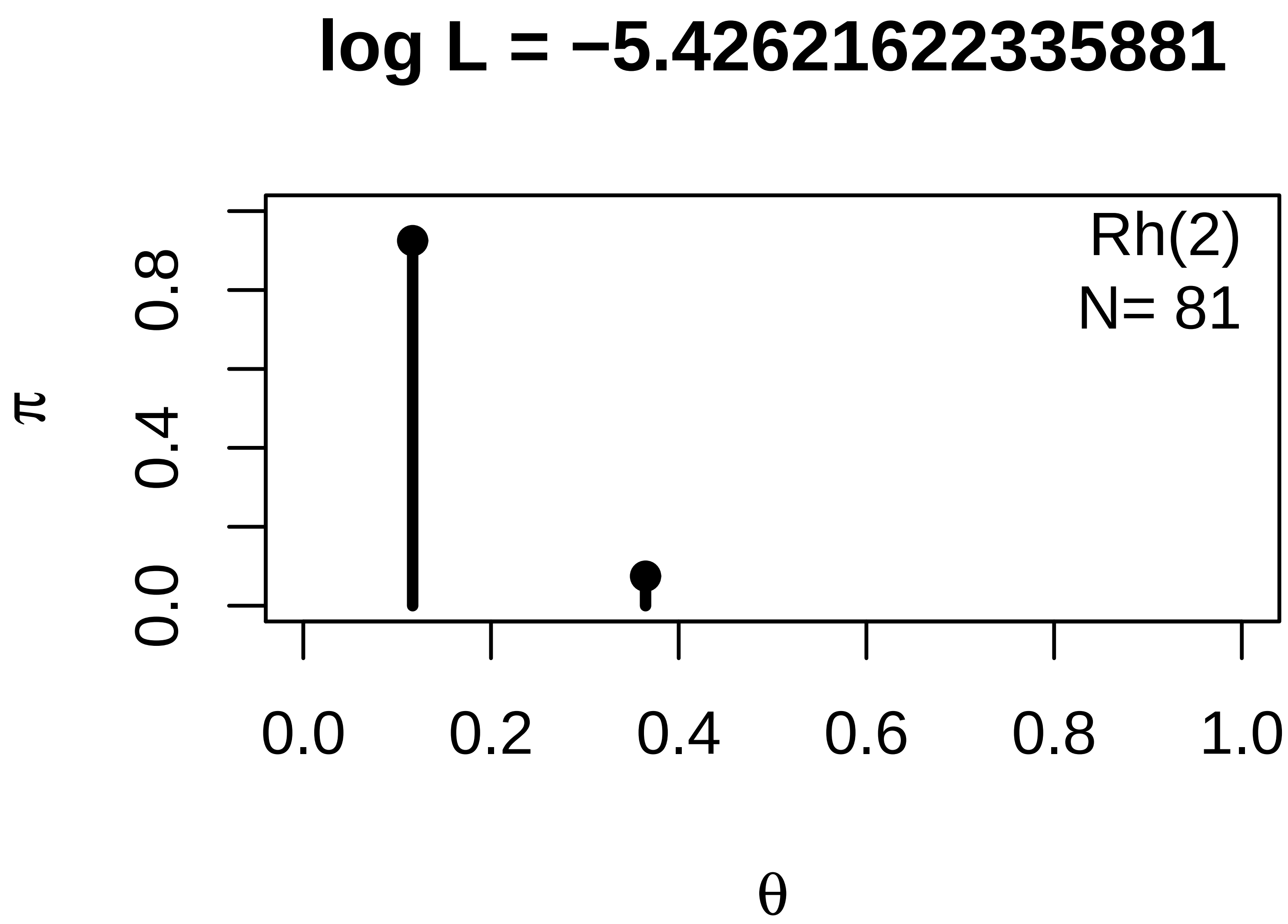
Step 88



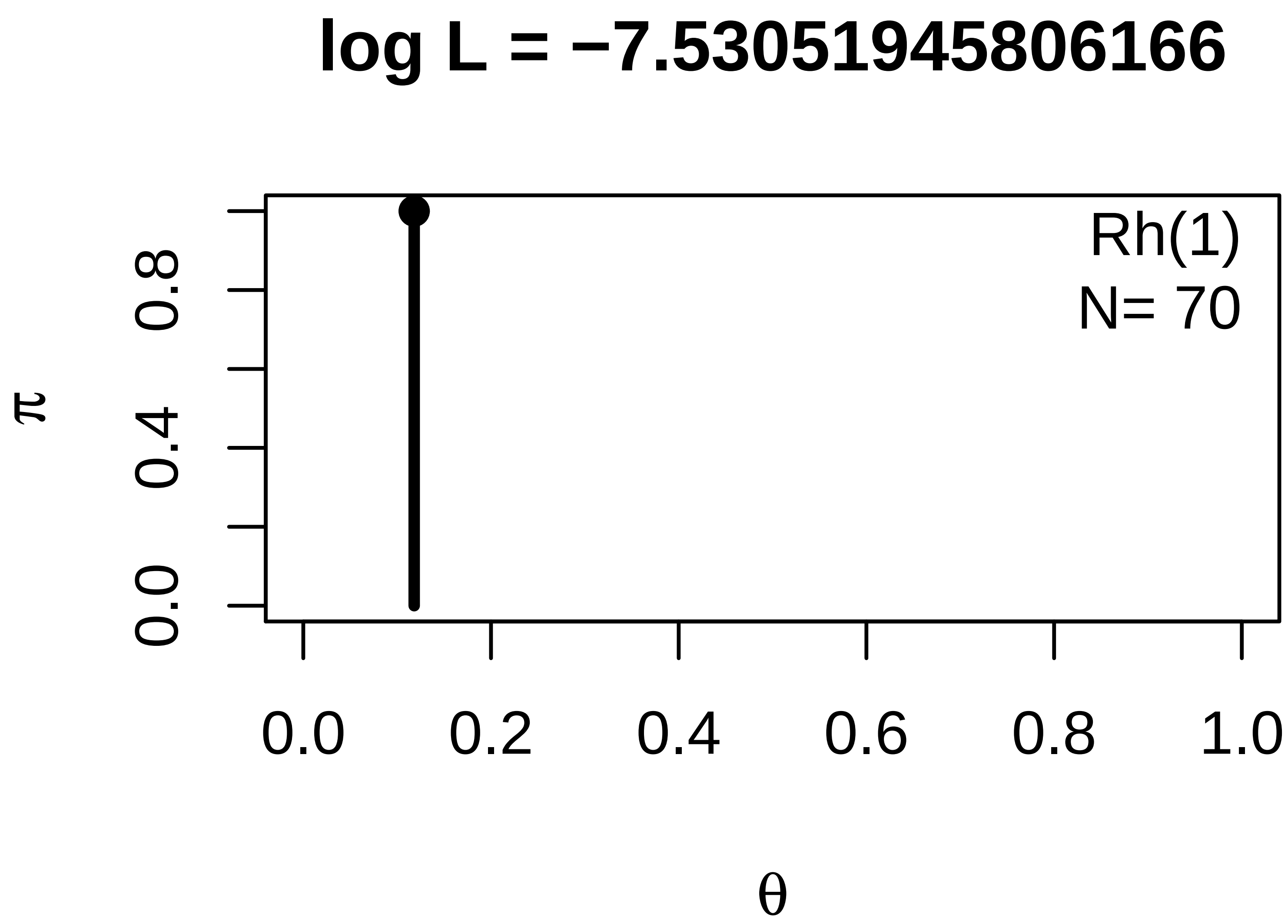
Step 89



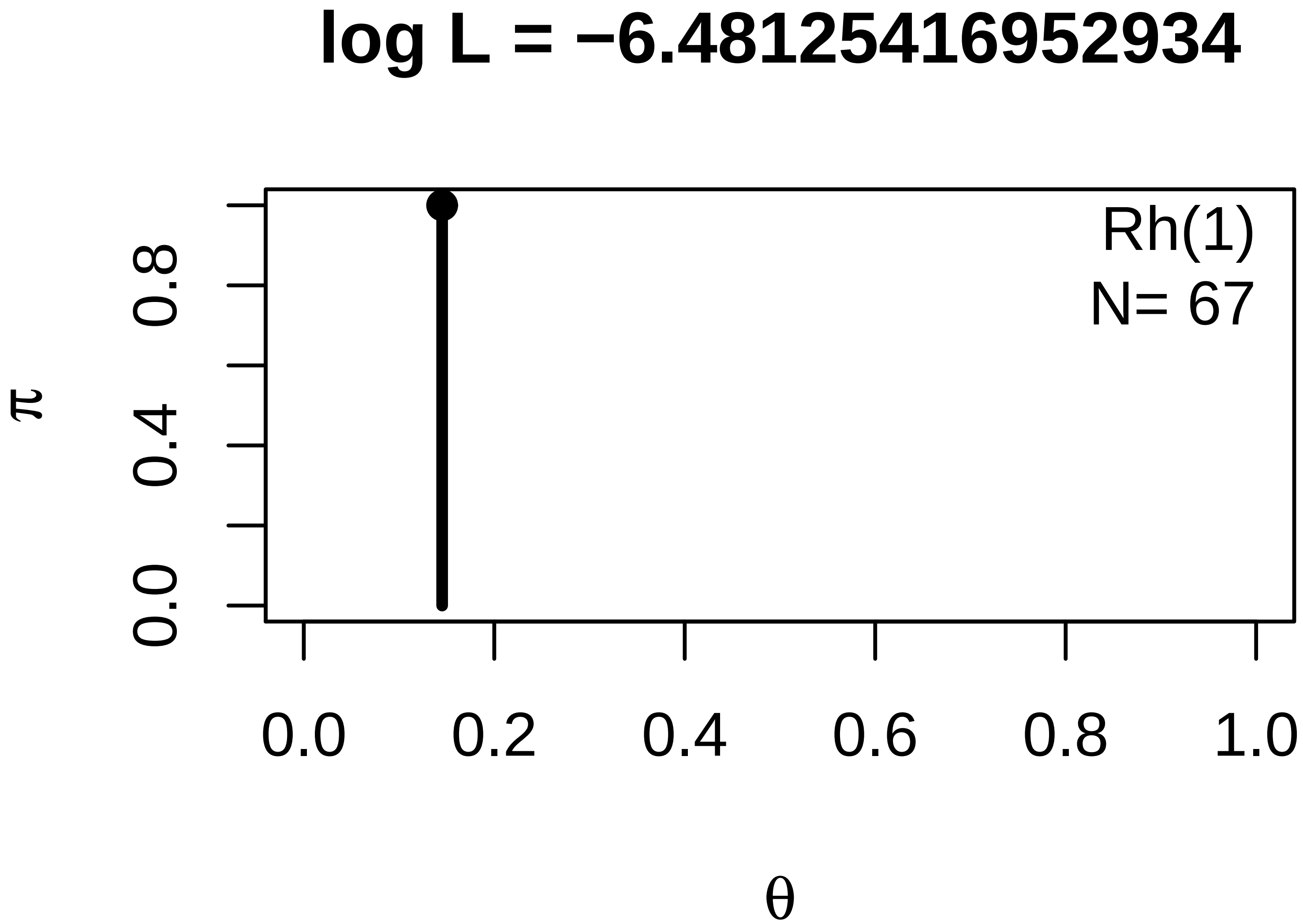
Step 90



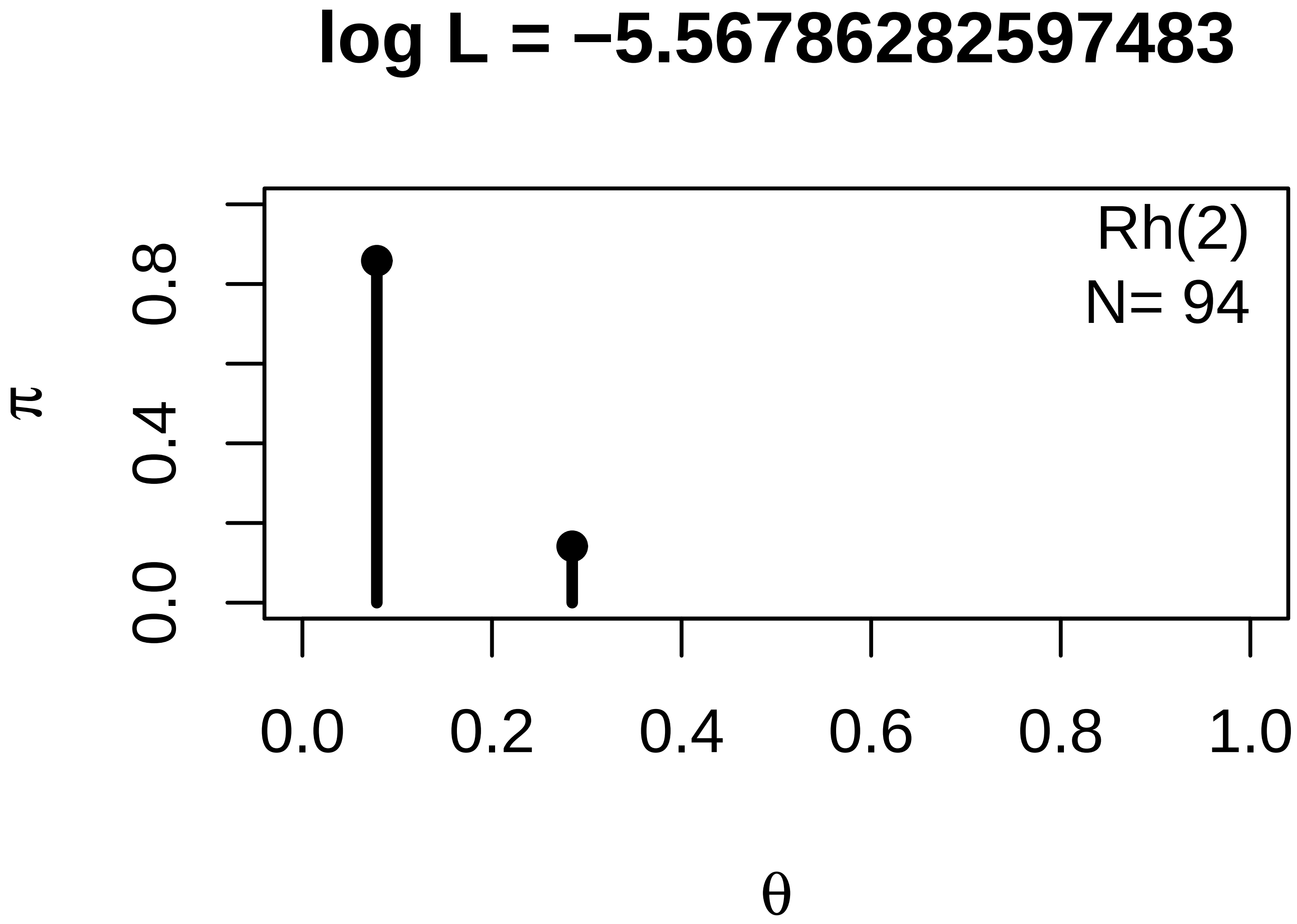
Step 91



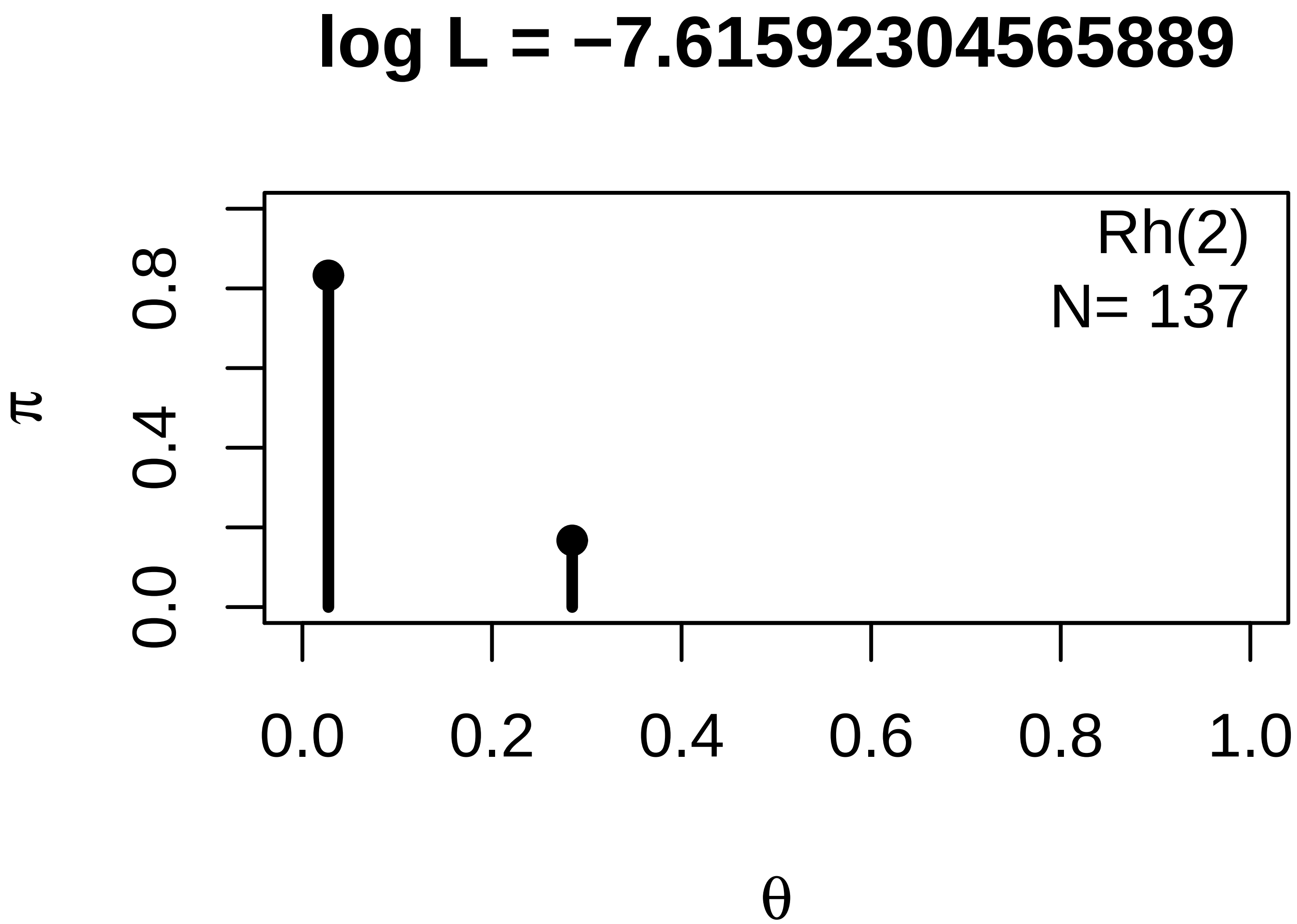
Step 92



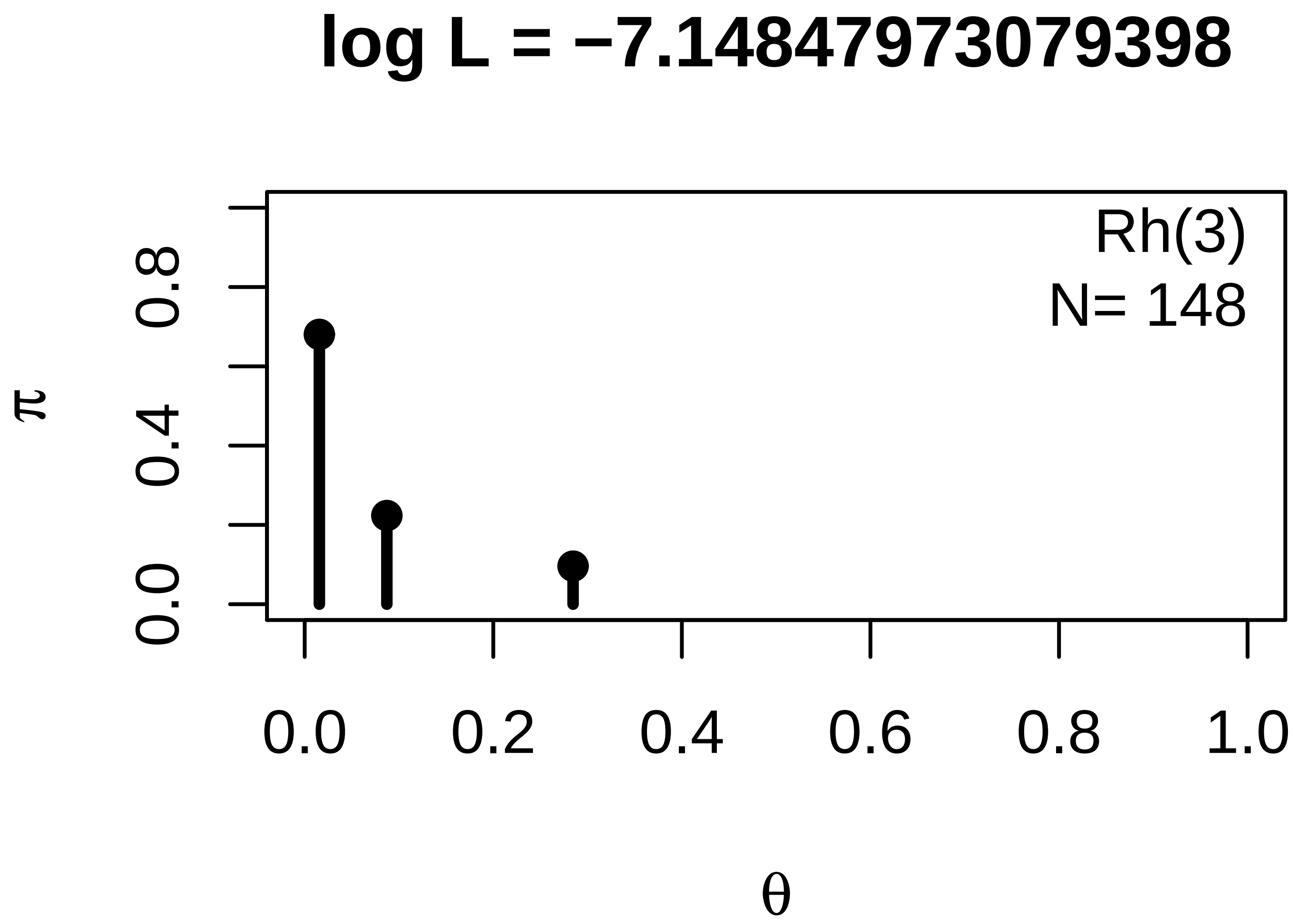
Step 93



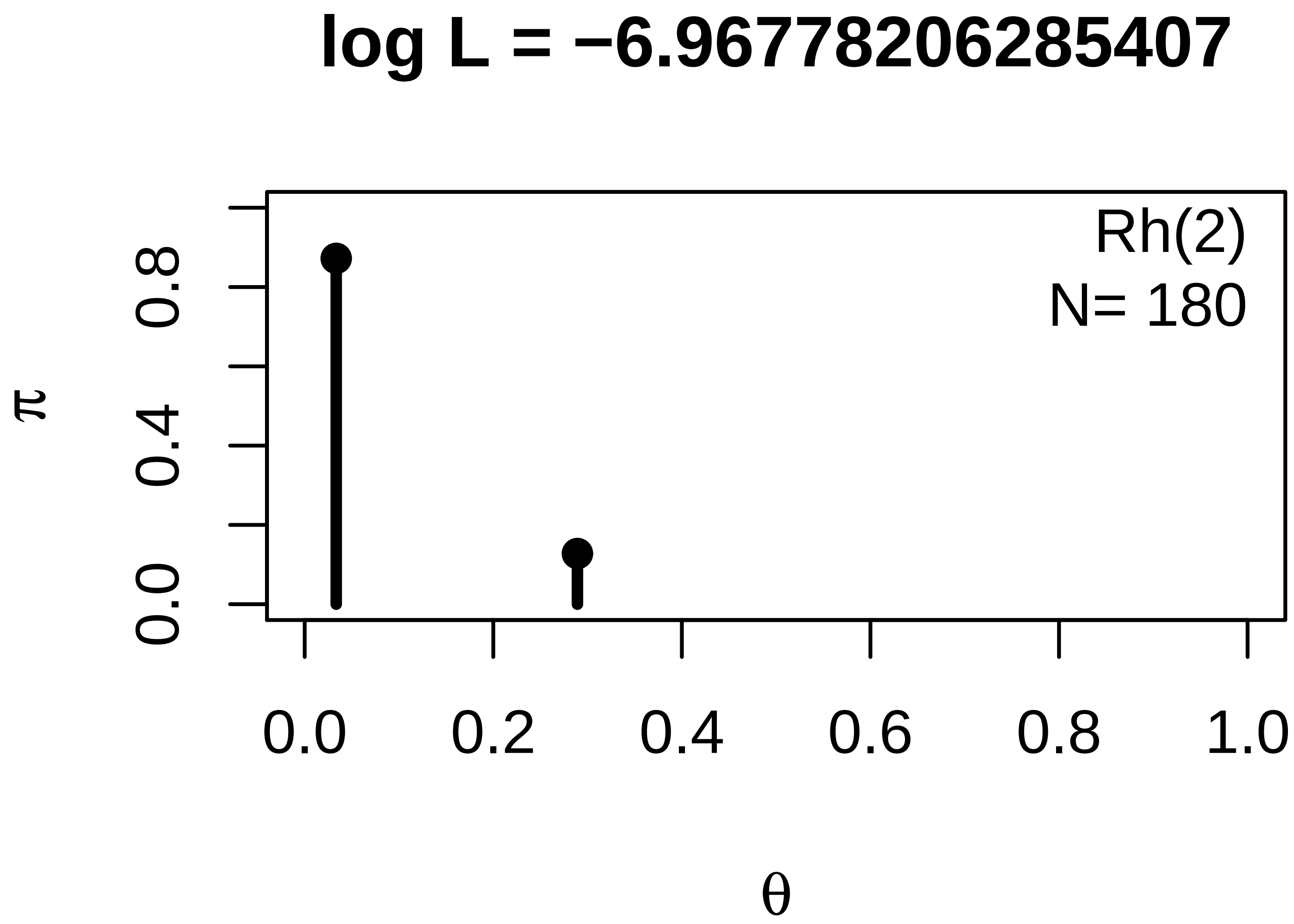
Step 94



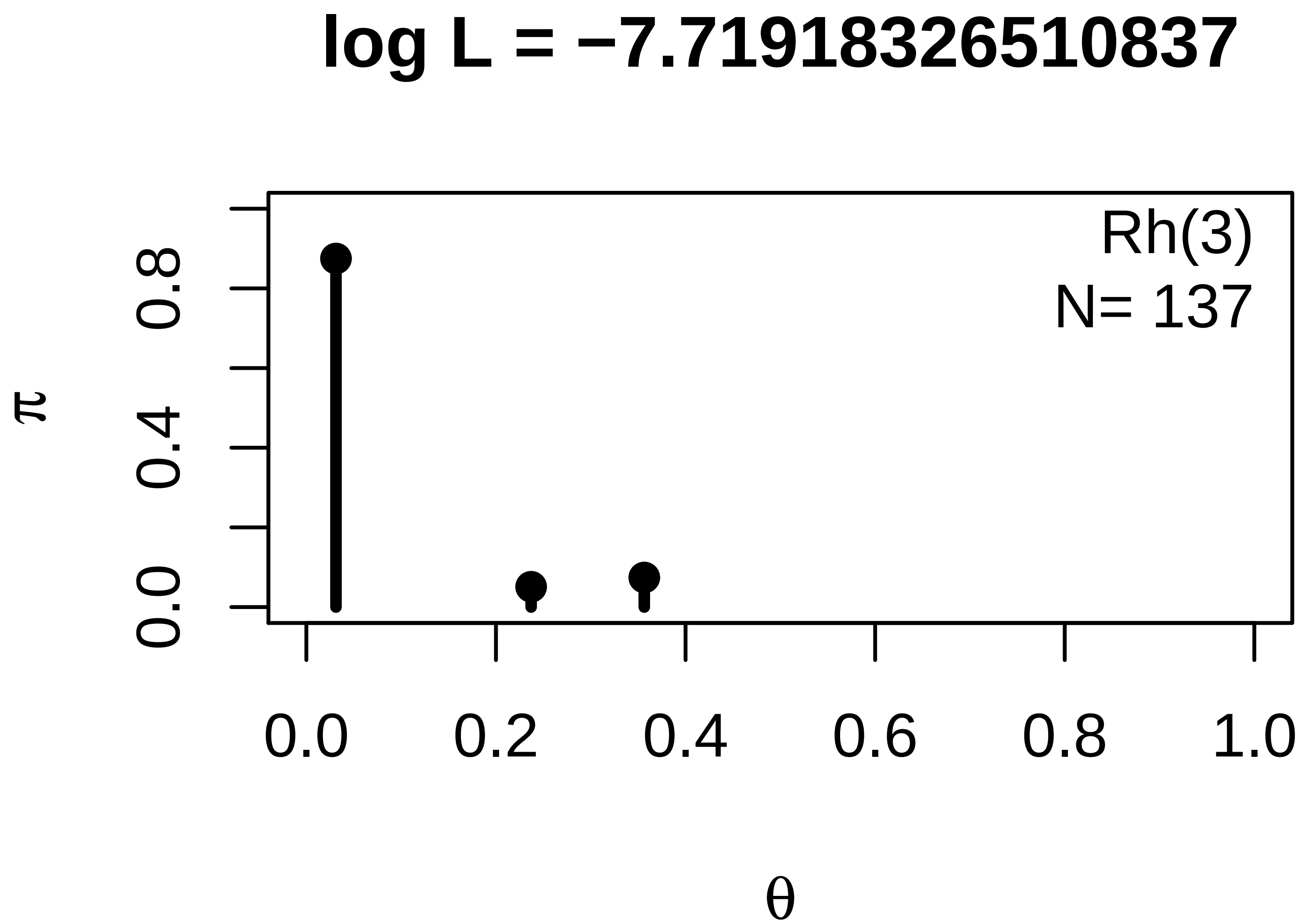
Step 95



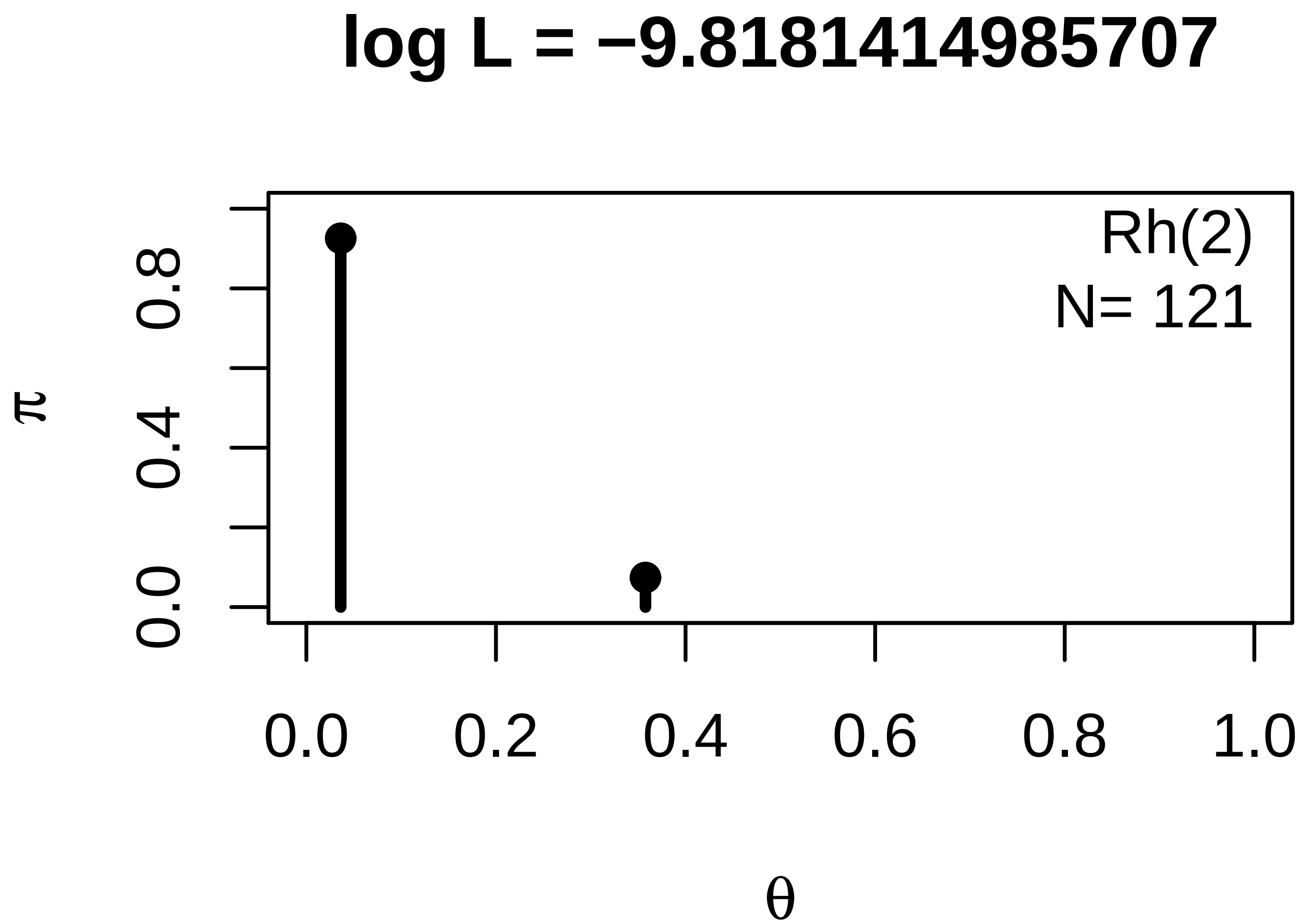
Step 96



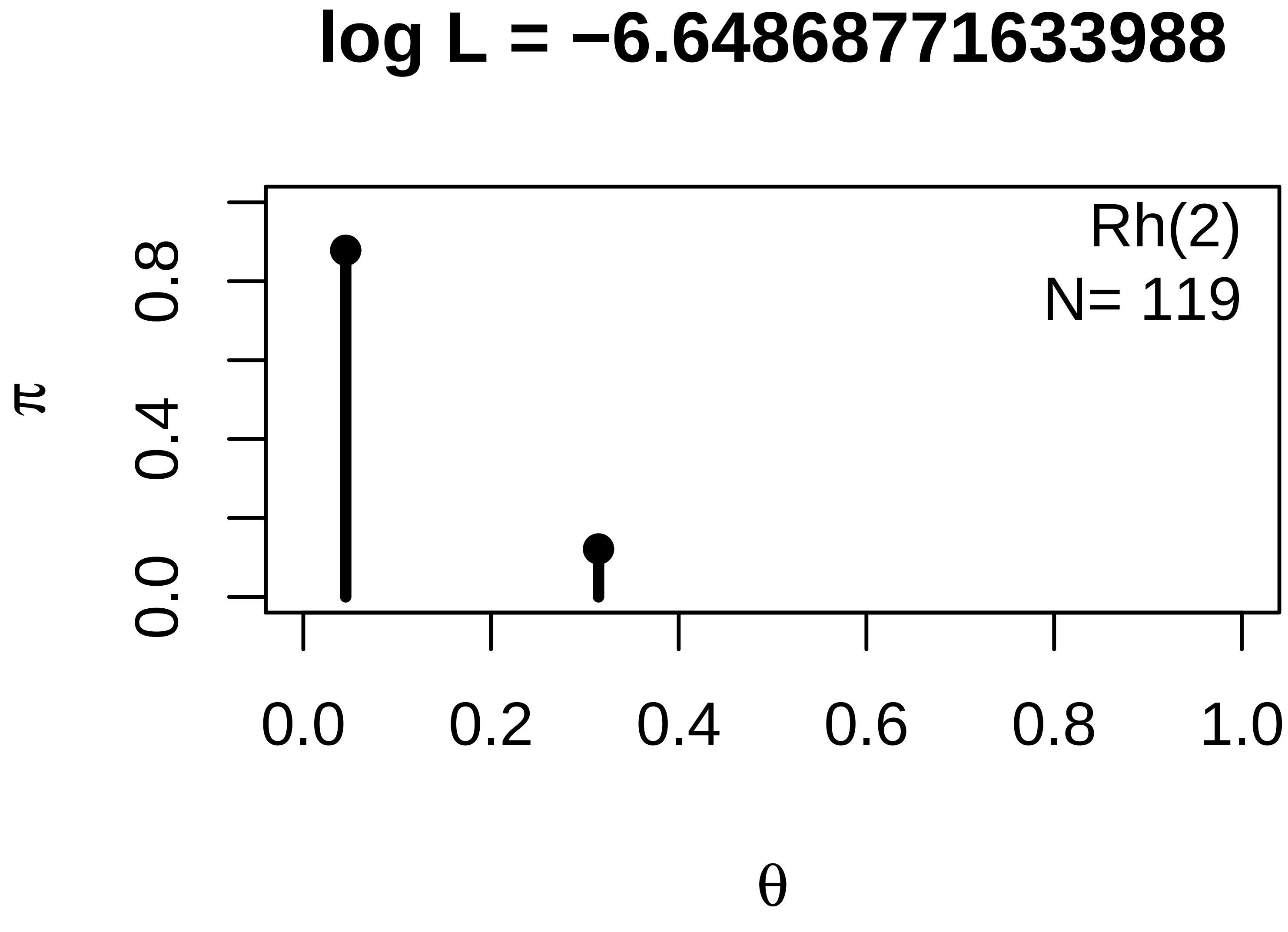
Step 97



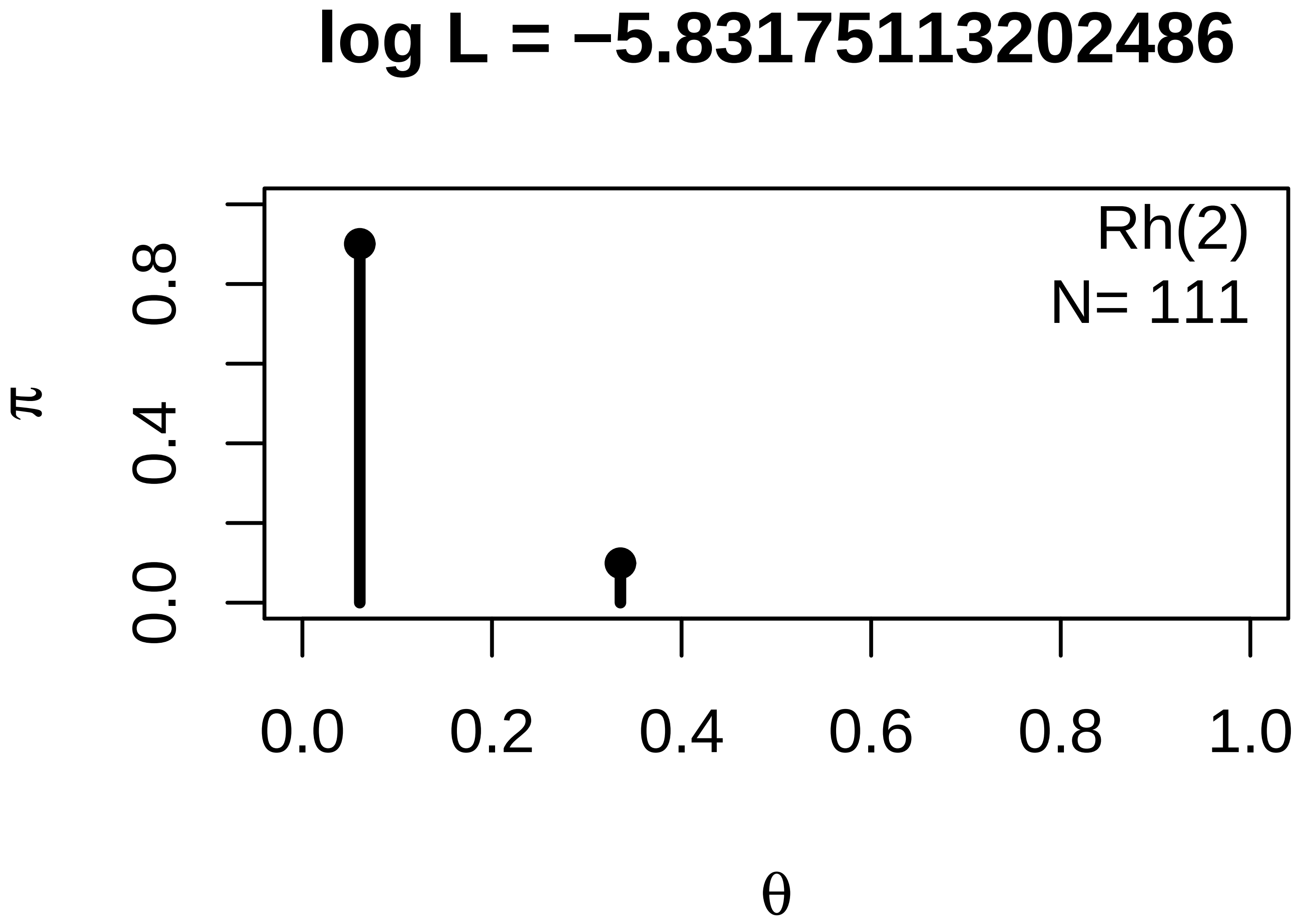
Step 98



Step 99

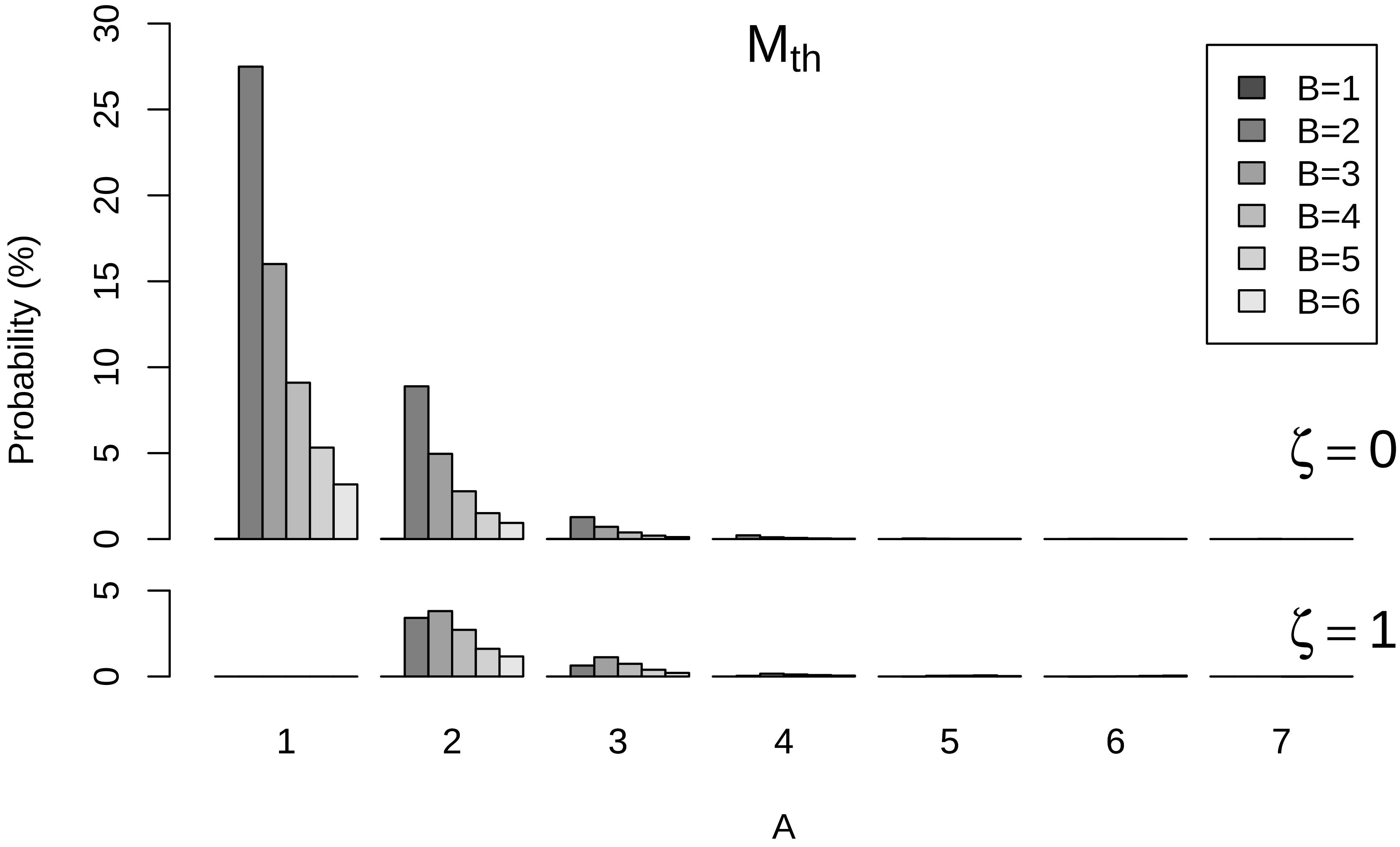


Step 100



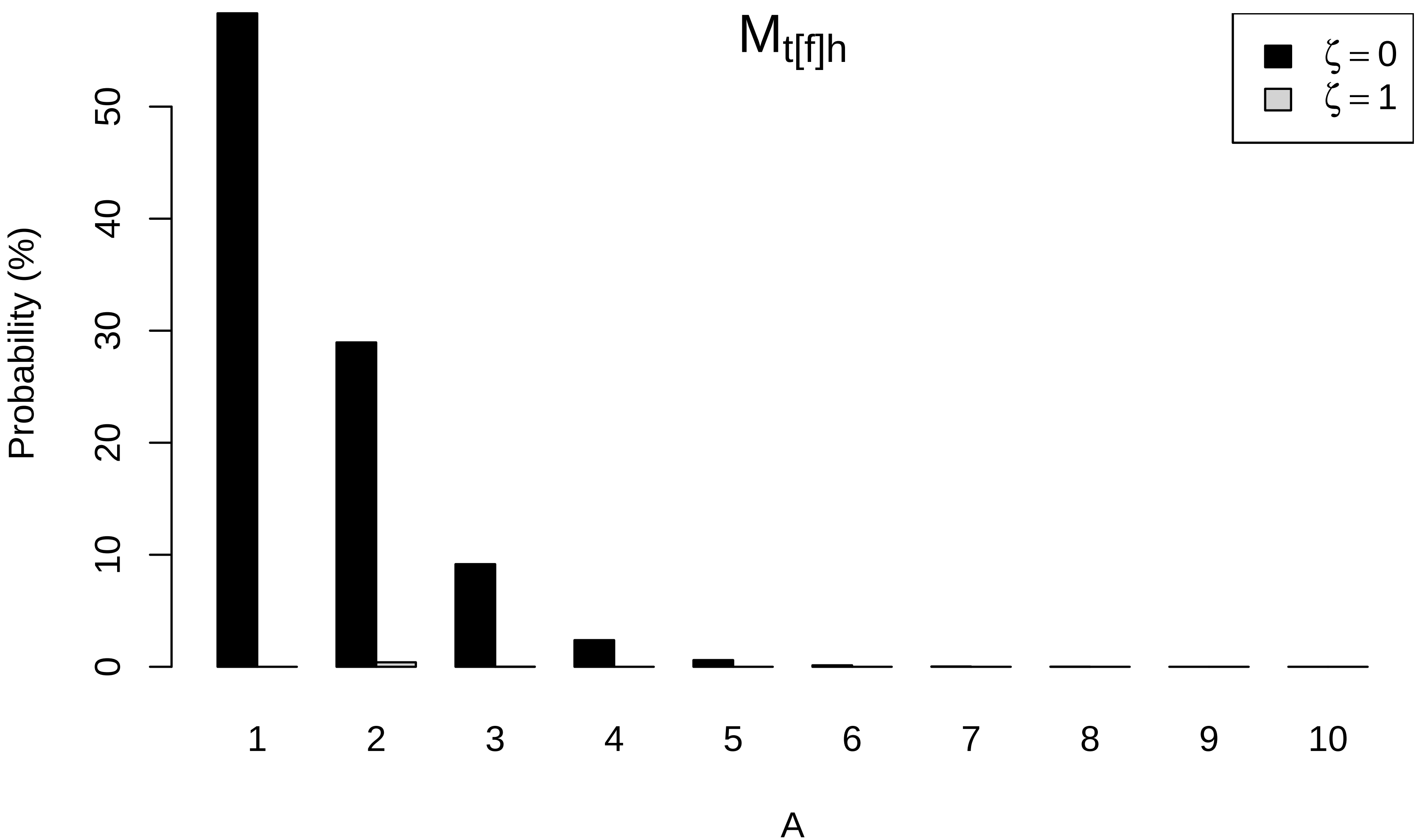
Application: AT&T Switch Testing

M_{th}: 2-way finite mixture



Application: AT&T Switch Testing

$M_{t[f]h}$: 1-way finite mixture with fixed sample effects



Posterior Model Probabilities

	$p(A,\zeta \mid X,M_{\text{t[f]h}})$		$p(A,B,\zeta \mid X,M_{\text{th}})$										
A	$\zeta=0$	$\zeta=1$	$\zeta=0$						$\zeta=1$				
$B=$			1	2	3	4	5	6	2	3	4	5	6
1	58.3			27.5	16.0	9.1	5.3	3.2					
2	29.0	0.4		8.9	5.0	2.8	1.5	0.9	3.4	3.8	2.7	1.6	1.2
3	9.2			1.3	0.7	0.4	0.2	0.1	0.6	1.1	0.7	0.4	0.2
4	2.4			0.2	0.1	0.1				0.2	0.1	0.1	0.1
5	0.6											0.1	
6	0.1												0.1
7													
8													
9													
10													

Posterior Means

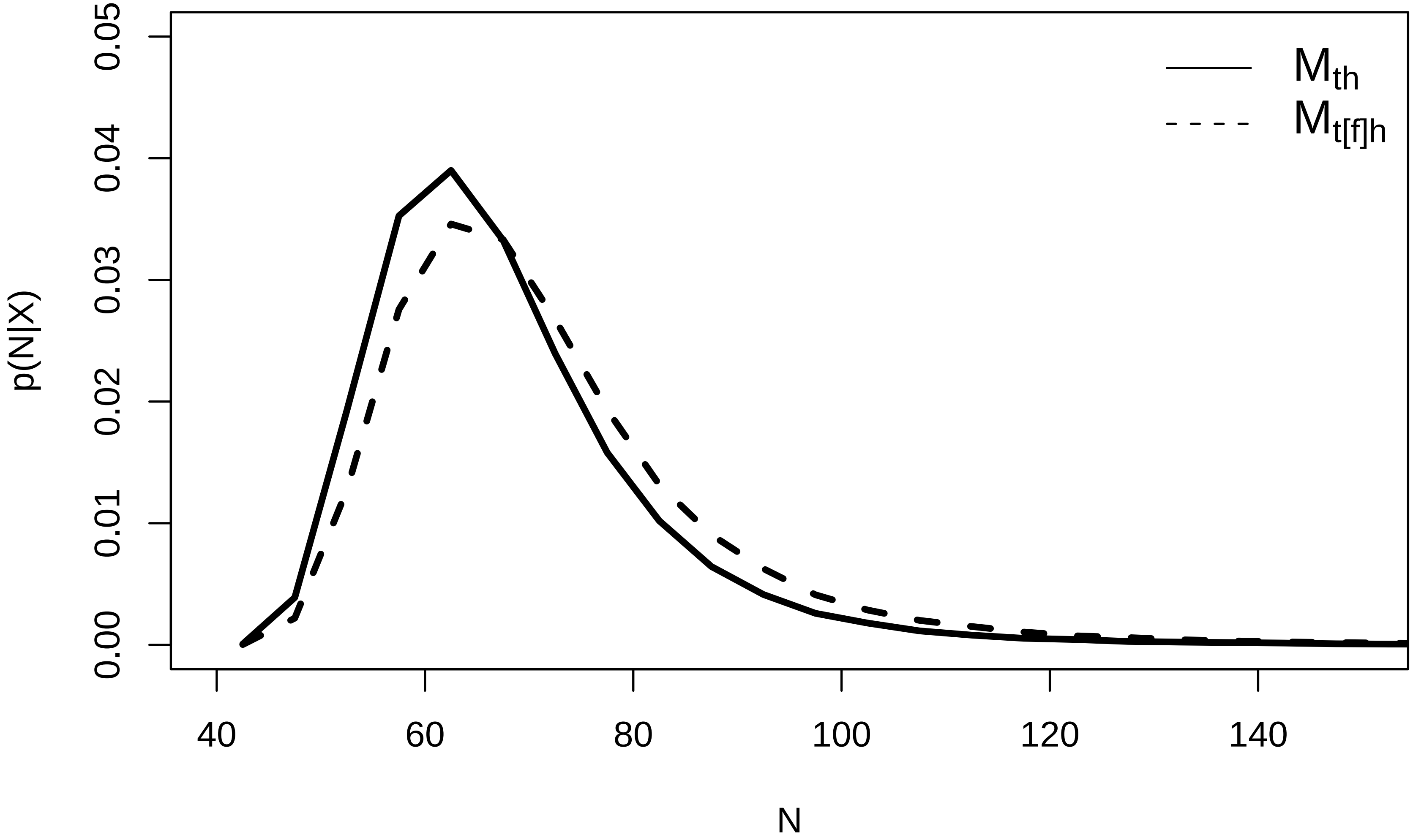
A	$E[N \mid X, M_{t[f]h}]$ $\zeta = 0$	$E[N \mid X, M_{t[f]h}]$ $\zeta = 1$	$E[N \mid X, M_{th}]$ $\zeta = 0$						$E[N \mid X, M_{th}]$ $\zeta = 1$				
1	67.8		78.5	65.9	64.7	64.3	64.2	64.2					
2	78.6	81.8	103.6	72.1	70.5	70.1	70.0	70.0	79.4	79.2	76.8	76.9	75.9
3	79.6	65.2	100.0	71.2	69.8	69.2	68.3	70.4	76.1	81.4	82.9	83.4	80.0
4	78.1			71.0	70.8	67.7	67.5	68.9	75.6	85.8	77.8	73.6	72.8
5	75.5			68.4	67.1	69.5	64.7	64.0	67.5	78.8	75.6	67.0	69.7
6	75.0			70.2	64.9	62.0	67.8	58.0	65.0	91.3	78.2	64.5	66.1
7	76.4				62.0						74.0	66.3	67.7
8	84.0												
9	83.0												
10	64.0												

Application: AT&T Switch Testing

Population size estimates $\widehat{N}$ in fixed and variable dimension models.

Model	Mean	Med.	95% CI	Npar
M_0	78.5	76	(63,101)	2
M_{t_2}	65.9	66	(54,97)	10
M_{t_3}	64.7	64	(55,77)	12
M_{t_4}	64.3	62	(60,64)	14
$M_{t_2+h_2}$	72.1	70	(63,77)	12
$M_{t_2\times h_2}$	79.4	73	(55,122)	13
M_{th}	68.7	66	(51,103)	
$M_{t[f]}$	67.8	66	(52,94)	8
$M_{t[f]+h_2}$	78.6	74	(54,129)	10
$M_{t[f]\times h_2}$				15
$M_{t[f]h}$	72.4	69	(52,114)	

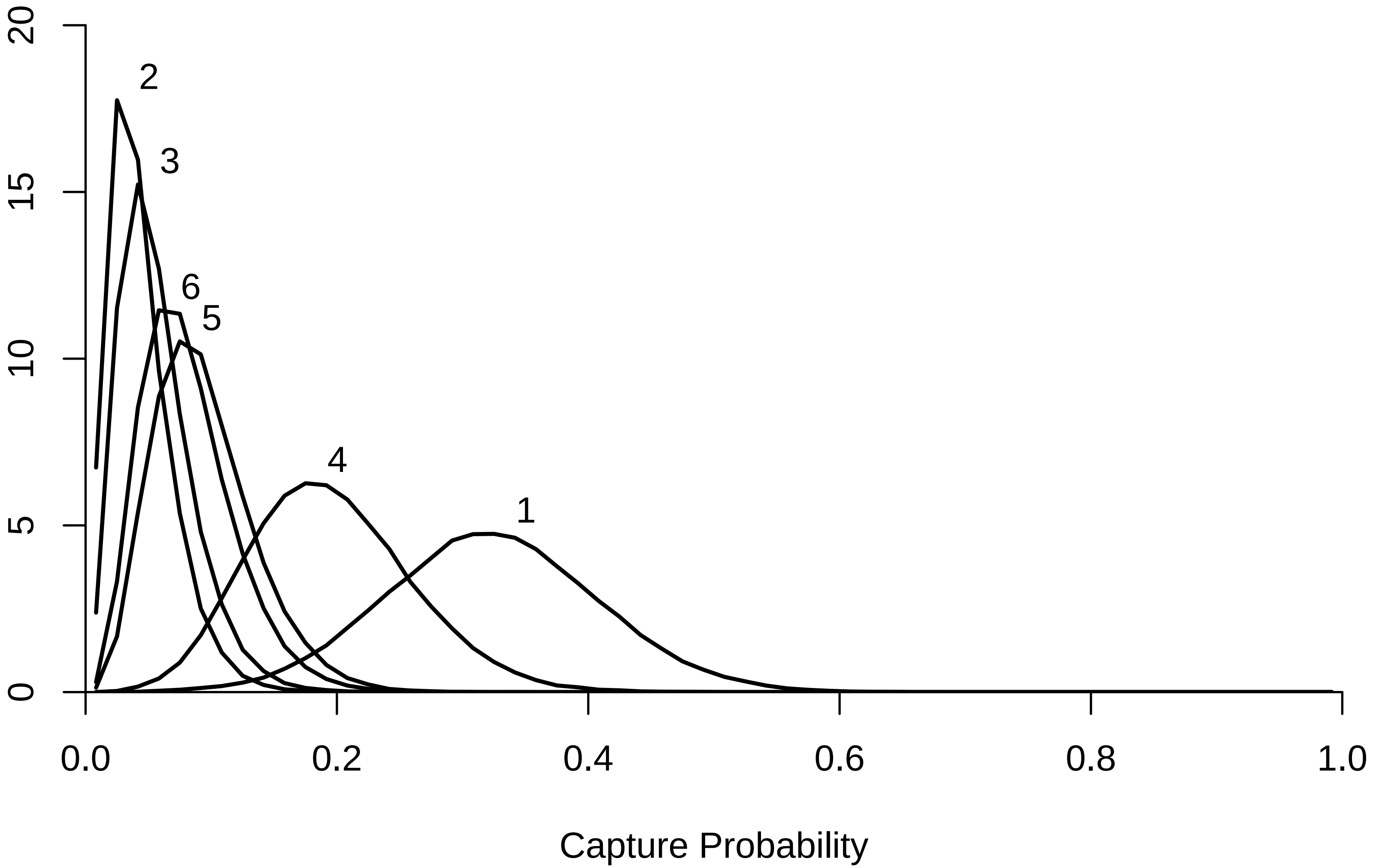
Posterior distribution of N



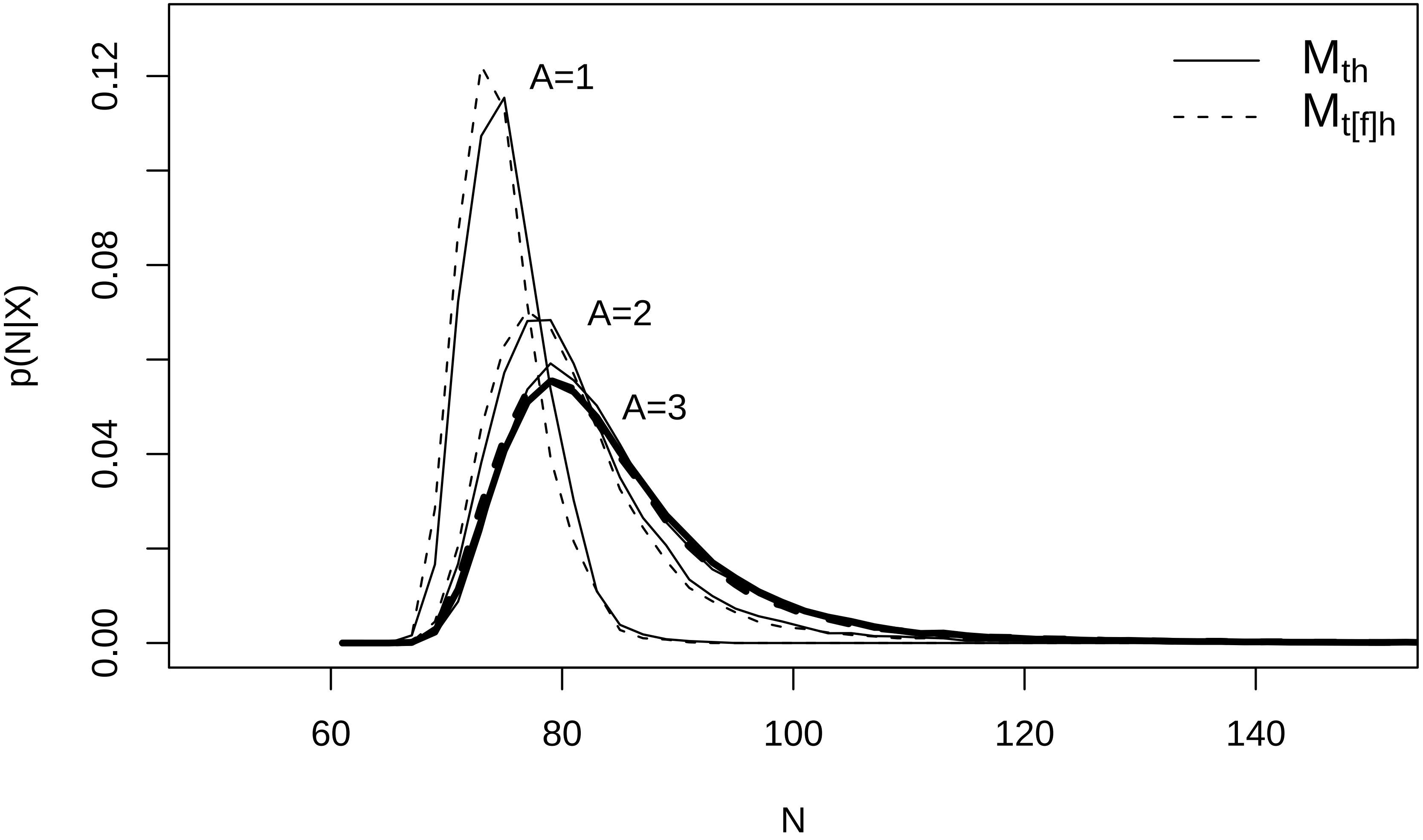
Application: AT&T Switch Testing

- Models M_{th} and $M_{\mathrm{t[f]h}}$ both predict of the order 70 errors in total (95% credible intervals [51,103] and [52,114])
- $\sim$ 30 errors not so far detected
- Evidence for heterogeneity among **reviewers**, but not amongst faults
- Bayes Factor $\mathrm{BF}(M_{\mathrm{th}} : M_{\mathrm{t[f]h}}) = 0.42$: neither model class convincingly favoured.

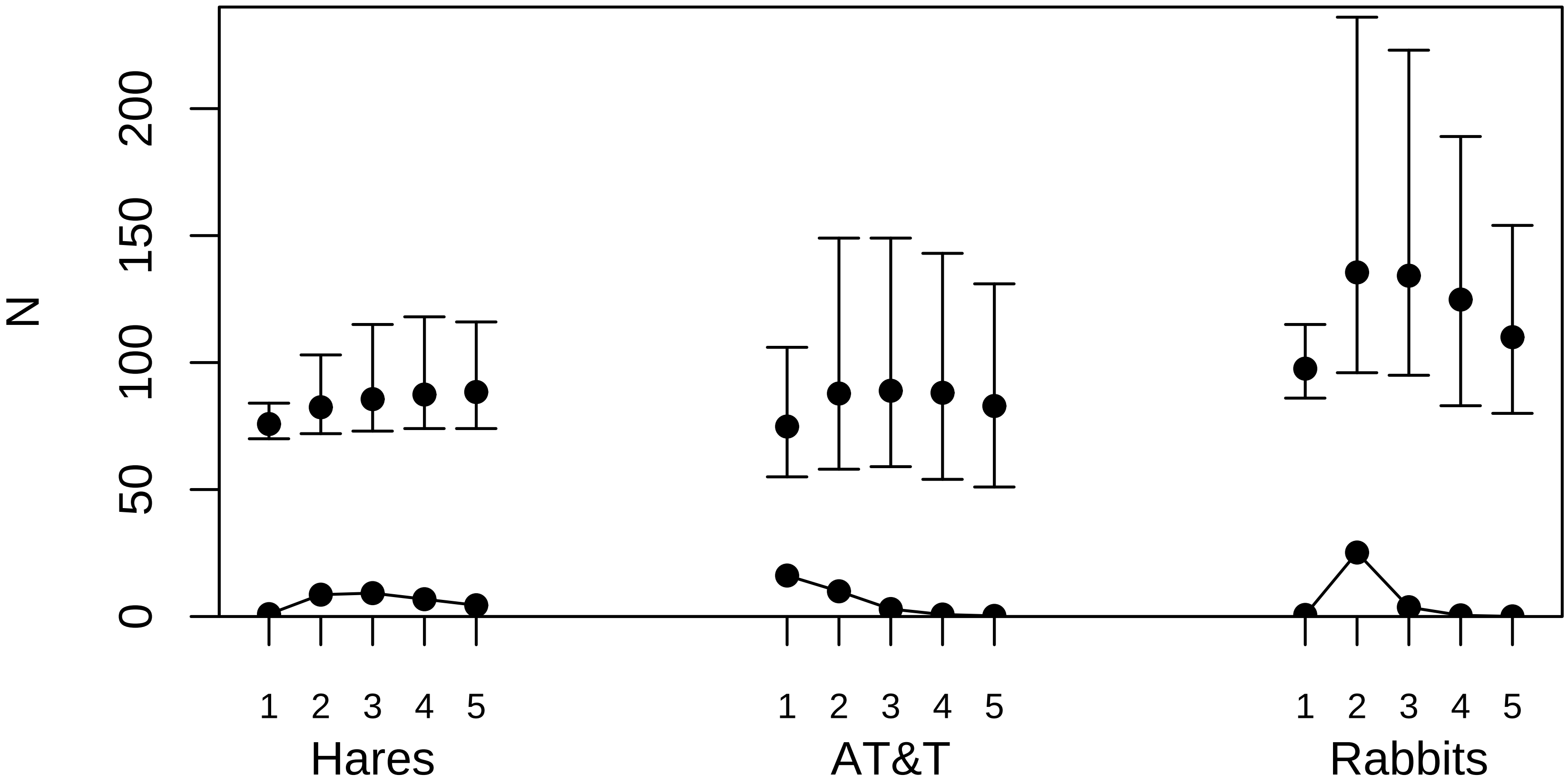
Skills of the 6 reviewers



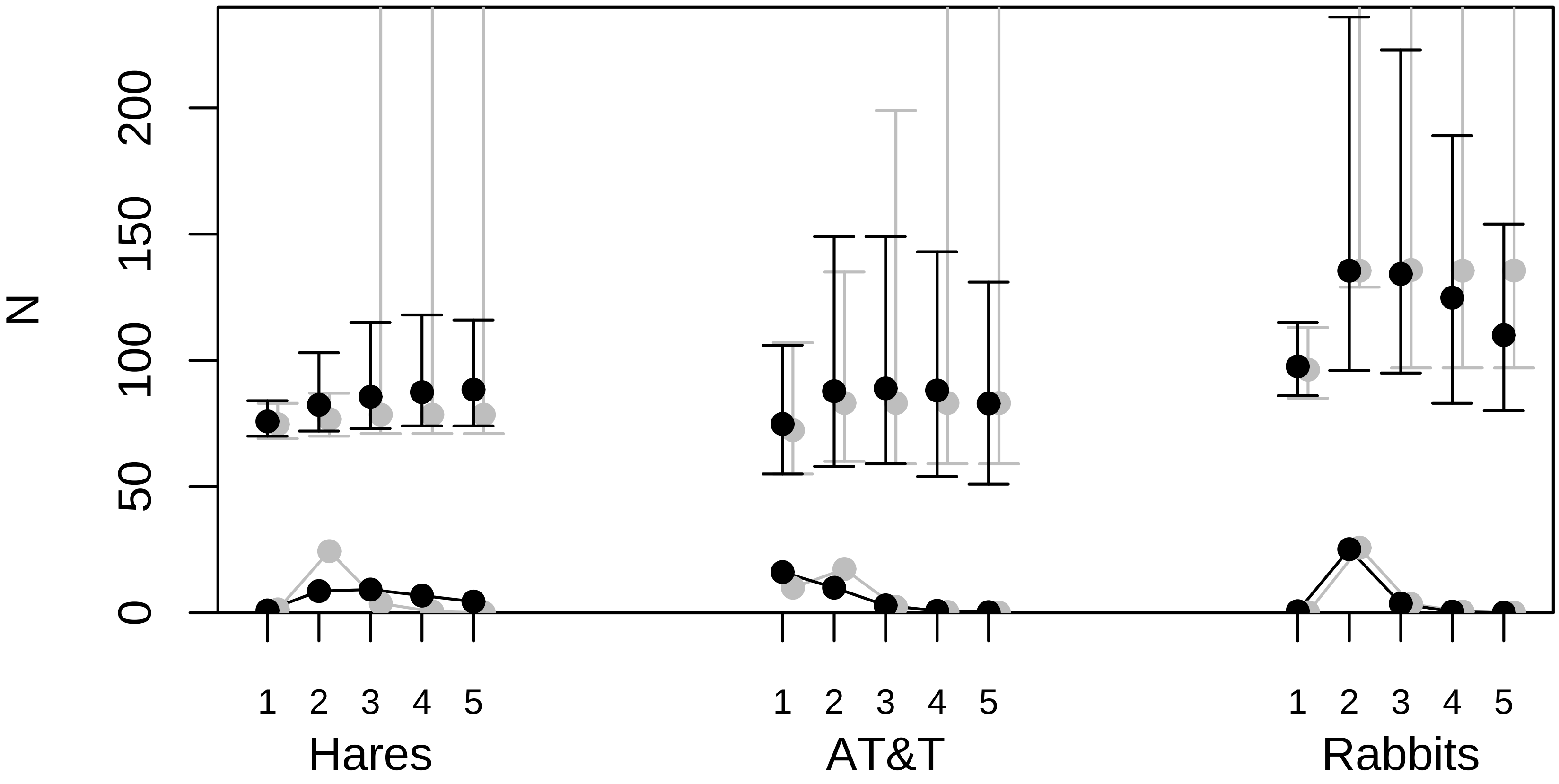
Application: Snowshoe Hares



Interval estimates – 3 examples



Interval estimates – 3 examples



Summary

- RJMCMC is a practicable means for model selection/averaging in Capture-Recapture with finite mixtures.
- Doesn't solve non-identifiability between model classes
- Priors regularise the likelihood – adding extra components does not affect estimates much.