The role of

Bias

in Statistics

D.A.S. Fraser
Statistics
U Toronto

April 12 2012
U Western Ontario

www.utstat.toronto.edu/dfraser/documents/uwo12.pdf
Joint work; references at end .../xxx.pdf

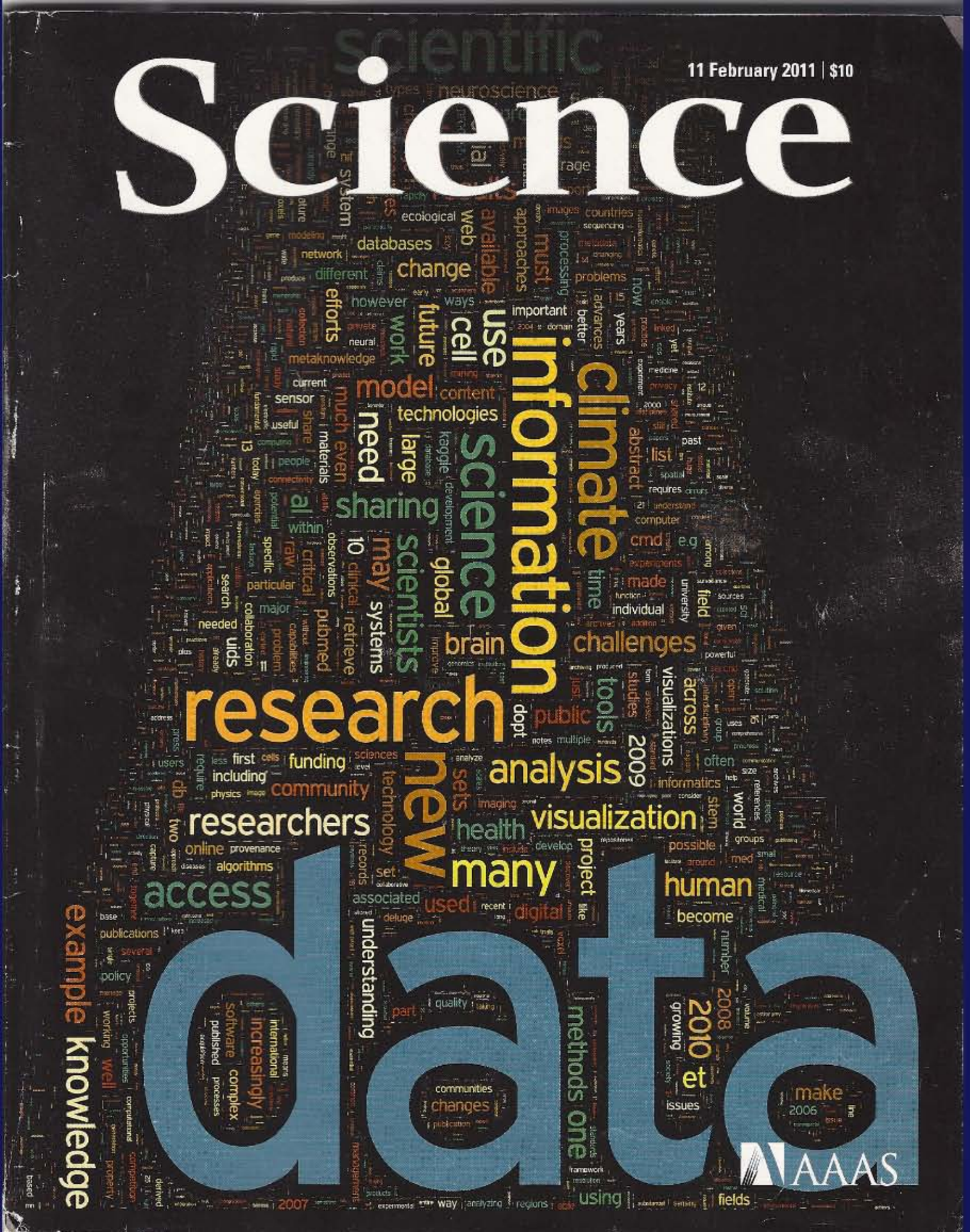
The journal Science recognizes that

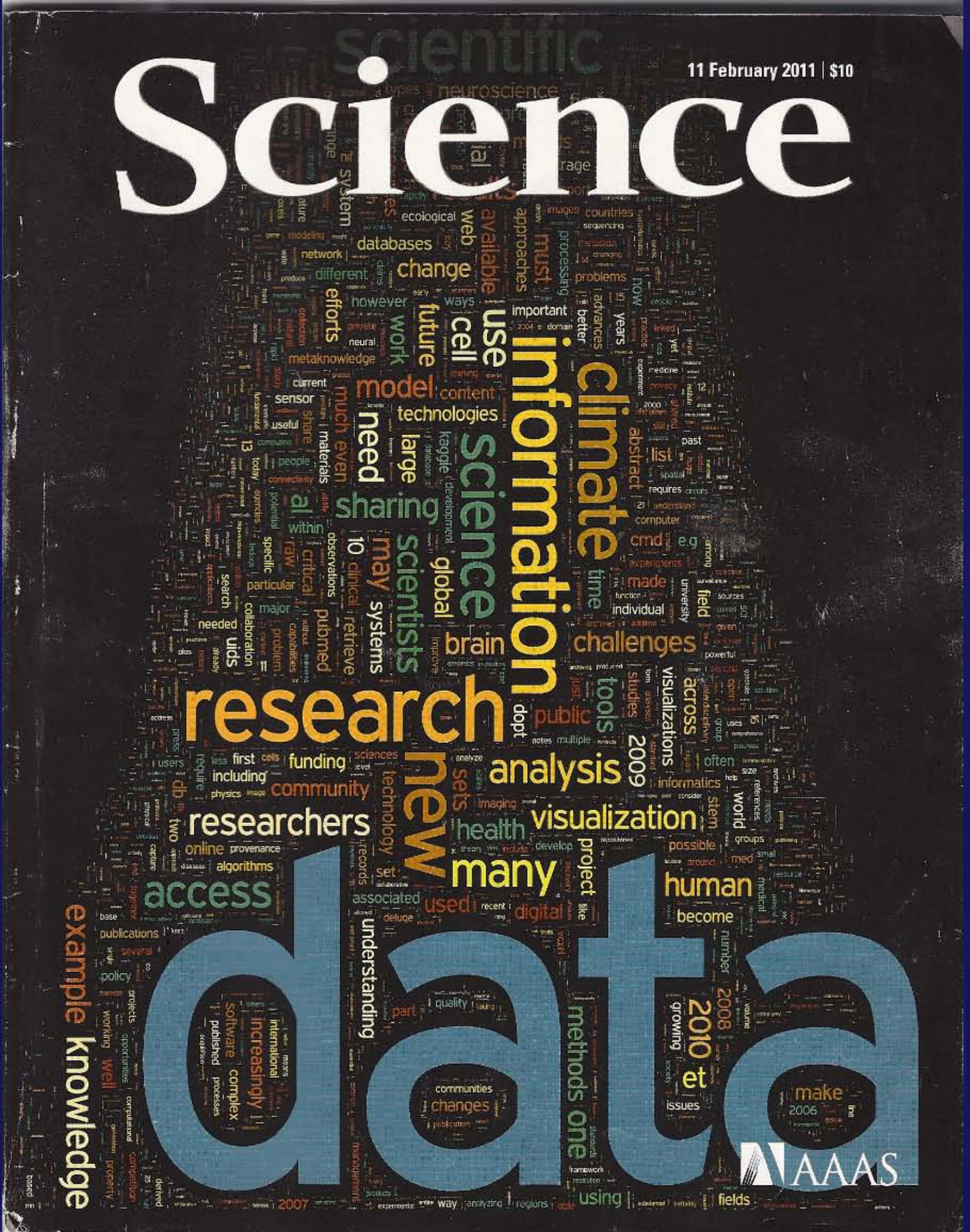
The journal **Science** recognizes that

Data are everywhere!

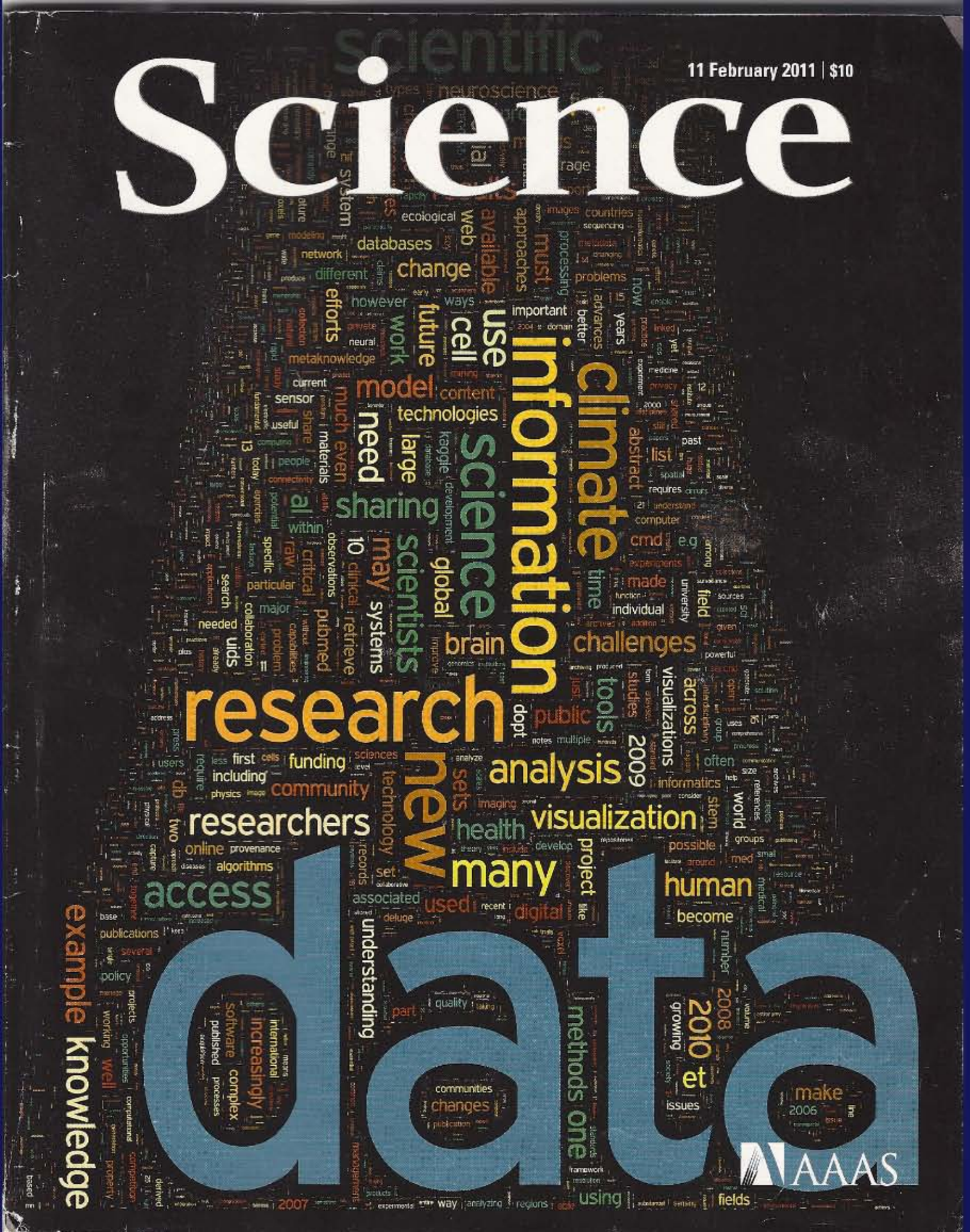
Science

11 February 2011 | $10



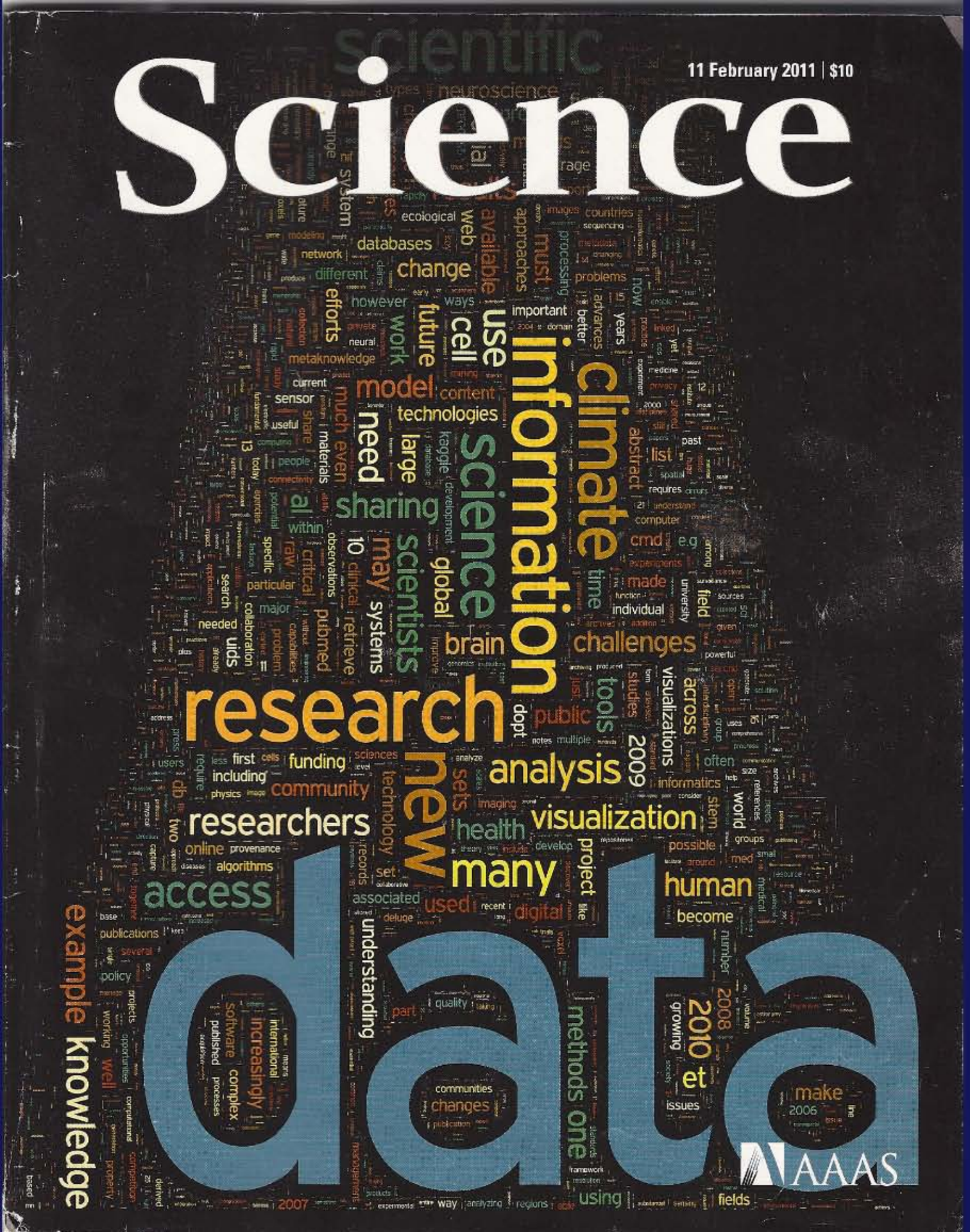


Science, Nature -
major science journals



Science, Nature -
major science journals

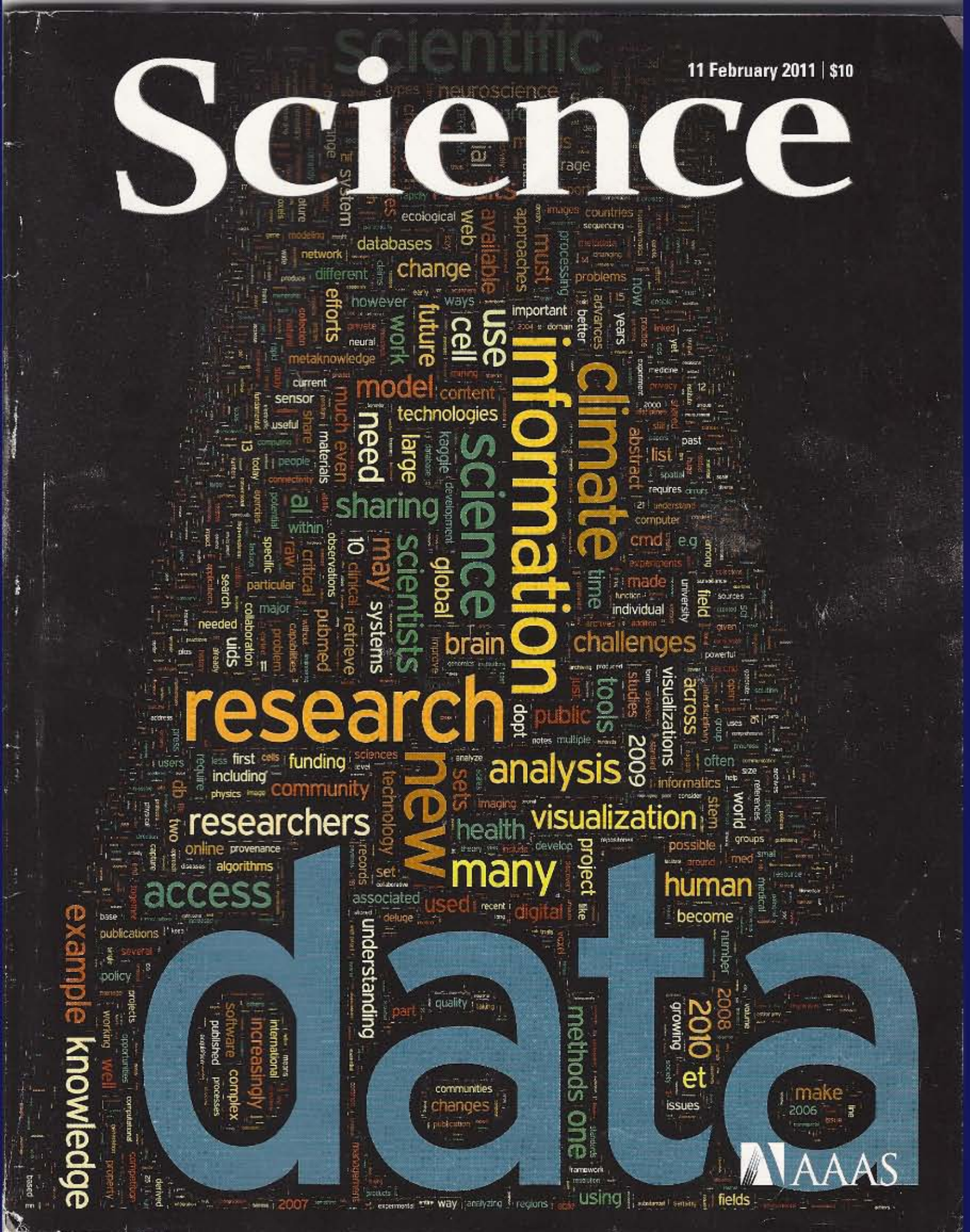
Issue - 38 pages on
Data
- 15 articles



Science, Nature -
major science journals

Issue - 38 pages on
Data
- 15 articles

See: Word cloud
Word histogram
- frequency of each
word in articles

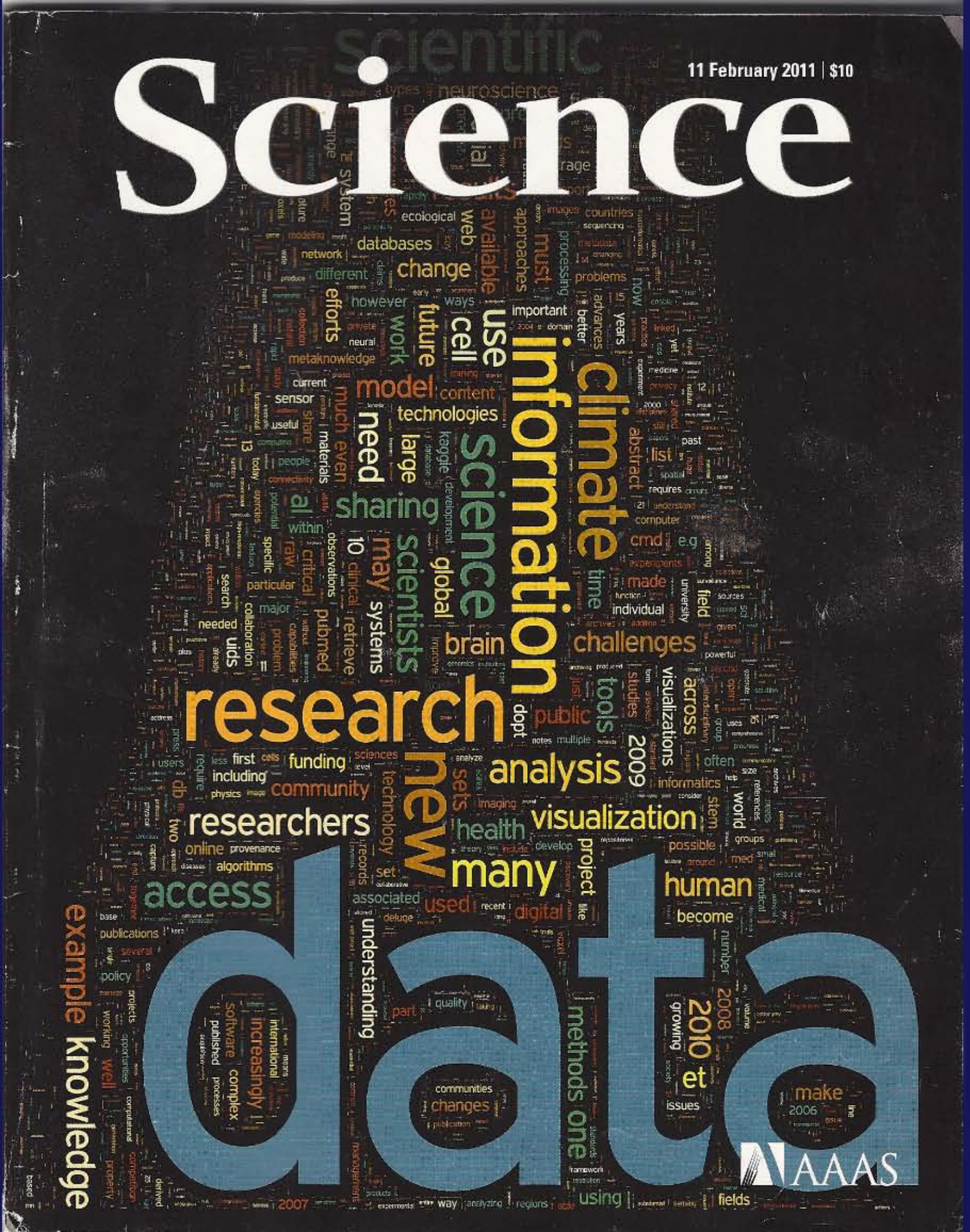


Science, Nature -
major science journals

Issue - 38 pages on
Data
- 15 articles

See: Word cloud
Word histogram
- frequency of each
word in articles

But: "Statistics"
is not there!



Science, Nature -
major science journals

Issue - 38 pages on
Data
- 15 articles

See: Word cloud
Word histogram
- frequency of each
word in articles

But: "Statistics"
is not there!

"Statistics" is not
part of Data!

Outline

- Data and Statistics:
  - (a) Journals: Data, Groves, Vioxx, Replication
- Statistics:
  - (b) Two logics
  - (c) What are they?
- Can Right prior do it all:
  - (d) Linear: Yes
  - (e) Curved: No
- How to get the "Right" prior
  - (f) from Continuity!
  - (g) from information!
- Measure bias in prior...(if not "Right")
  - (h) Available to 2nd order

Andrew Grove is the former Chief Executive Officer of Intel Corporation. He is a patient advocate at the University of California, San Francisco.

Rethinking Clinical Trials

THE BIOMEDICAL INDUSTRY SPENDS OVER $50 BILLION PER YEAR ON RESEARCH AND DEVELOPMENT and produces some 20 new drugs. One reason for this disappointing output is the byzantine U.S. clinical trial system that requires large numbers of patients. Half of all trials are delayed, 80 to 90% of them because of a shortage of trial participants. Patient limitations also cause large and unpredicted expenses to pharmaceutical and biotech companies as they are forced to tread water. As the industry moves toward biologics and personalized medicine, this limitation will become even greater. A breakthrough in regulation is needed to create a system that does more with fewer patients.

The current clinical trial system in the United States is more than 50 years old. Its architecture was conceived when electronic manipulation of data was limited, slow, and expensive. Since then, network and connectivity costs have declined ten thousand–fold, data storage costs over a million-fold, and computation costs by an even larger factor. Today, complex and powerful applications like electronic commerce are deployed on a large scale. Amazon.com is a good example. A large database of customers and products form the kernel of its operation. A customer's characteristics (like buying history and preferences) are observed and stored. Customers can be grouped and the buying behavior of any individual or group can be compared with corresponding behavior of others. Amazon can also track how a group or an individual responds to an outside action (such as advertising).

We might conceptualize an "e-trial" system along similar lines. Drug safety would continue to be ensured by the U.S. Food and Drug Administration. While safety-focused Phase I trials would continue under their jurisdiction, establishing efficacy would no longer be under their purview. Once safety is proven, patients could access the medicine in question through qualified physicians. Patients' responses to a drug would be stored in a database, along with their medical histories. Patient identity would be protected by biometric identifiers, and the database would be open to qualified medical researchers as a "commons." The response of any patient or group of patients to a drug or treatment would be tracked and compared to those of others in the database who were treated in a different manner or not at all. These comparisons would provide insights into the factors that determine real-life efficacy: how individuals or subgroups respond to the drug. This would liberate drugs from the tyranny of the averages that characterize trial information today. The technology would facilitate such comparisons at incredible speeds and could quickly highlight negative results. As the patient population in the database grows and time passes, analysis of the data would also provide the information needed to conduct postmarketing studies and comparative effectiveness research.

Today's e-commerce systems started small and took nearly 20 years to develop. Adapting this kind of capability to medical information would be a monumental undertaking. Initiating and overseeing it would be an appropriate task for the professional societies. There are encouraging signs, including a call in 2004 by the American Medical Association for public registries of drugs, as well as a proposal for trials that incorporate feed-forward mechanisms.* Another proposal would allow patients to choose between medicines whose efficacy has been determined in different manners.† There is also a suggestion to use control of pricing to encourage drug developers to move forward in a "progressive" trial design.‡ Ideas, however, are not enough. We need the professions to mobilize and take advantage of this enormous opportunity.

– Andrew Grove

10.1126/science.1212118

*H. Hvitfeldt *et al.*, *Qual. Manag. Health Care* **18**, 247 (2009). †B. J. Madden, *Free to Choose Medicine* (Heartland Institute, Chicago, IL, 2010). ‡M. Boldrin, S. J. Swamidass, "A New Bargain for Drug Approvals," *Wall Street Journal*, 25 July 2011.

Another issue of Science (Sep 23/11)

Grove --- former CEO Intel

Andrew Grove is the former Chief Executive Officer of Intel Corporation. He is a patient advocate at the University of California, San Francisco.

Rethinking Clinical Trials

THE BIOMEDICAL INDUSTRY SPENDS OVER $50 BILLION PER YEAR ON RESEARCH AND DEVELOPMENT and produces some 20 new drugs. One reason for this disappointing output is the byzantine U.S. clinical trial system that requires large numbers of patients. Half of all trials are delayed, 80 to 90% of them because of a shortage of trial participants. Patient limitations also cause large and unpredicted expenses to pharmaceutical and biotech companies as they are forced to tread water. As the industry moves toward biologics and personalized medicine, this limitation will become even greater. A breakthrough in regulation is needed to create a system that does more with fewer patients.

The current clinical trial system in the United States is more than 50 years old. Its architecture was conceived when electronic manipulation of data was limited, slow, and expensive. Since then, network and connectivity costs have declined ten thousand–fold, data storage costs over a million-fold, and computation costs by an even larger factor. Today, complex and powerful applications like electronic commerce are deployed on a large scale. Amazon.com is a good example. A large database of customers and products form the kernel of its operation. A customer's characteristics (like buying history and preferences) are observed and stored. Customers can be grouped and the buying behavior of any individual or group can be compared with corresponding behavior of others. Amazon track how a group or an individual responds to an outsid (such as advertising).

under their purview. Once safety is proven, patients could access the medicine in question through qualified physicians. Patients' responses to a drug would be stored in a database, along with their medical histories. Patient identity would be protected by biometric identifiers, and the database would be open to qualified medical researchers as a "commons." The response of any patient or group of patients to a drug or treatment would be tracked and compared to those of others in the database who were treated in a different manner or not at all. These comparisons would provide insights into the factors that determine real-life efficacy: how individuals or subgroups respond to the drug. This would liberate drugs from the tyranny of the averages that characterize trial information today. The technology would facilitate such comparisons at incredible speeds and could quickly highlight negative results. As the patient population in the database grows and time passes, analysis of the data would also provide the information needed to conduct postmarketing studies and comparative effectiveness research.

Today's e-commerce systems started small and took nearly 20 years to develop. Adapting this kind of capability to medical information would be a monumental undertaking. Initiating and overseeing it would be an appropriate task for the professional societies. There are encouraging signs, including a call in 2004 by the American Medical Association for public registries of drugs, as well as a proposal for trials that incorporate feed-forward mechanisms.* Another proposal would allow patients to choose between medicines whose efficacy has been determined in different manners.† There is also a suggestion to use control of pricing to encourage drug developers to move forward in a "progressive" trial design.‡ Ideas, however, are not enough. We need the professions to mobilize and take advantage of this enormous opportunity.

– Andrew Grove

10.1126/science.1212118

*H. Hvitfeldt *et al.*, *Qual. Manag. Health Care* **18**, 247 (2009). †B. J. Madden, *Free to Choose Medicine* (Heartland Institute, Chicago, IL, 2010). ‡M. Boldrin, S. J. Swamidass, "A New Bargain for Drug Approvals," *Wall Street Journal*, 25 July 2011.

Another issue of Science Sep 23/11

Grove --- former CEO Intel

Once safety is proven, patients could access the medicine in question through qualified physicians.

⇒ immense database for medical researchers!

Andrew Grove is the former Chief Executive Officer of Intel Corporation. He is a patient advocate at the University of California, San Francisco.

Rethinking Clinical Trials

THE BIOMEDICAL INDUSTRY SPENDS OVER $50 BILLION PER YEAR ON RESEARCH AND DEVELOPMENT and produces some 20 new drugs. One reason for this disappointing output is the byzantine U.S. clinical trial system that requires large numbers of patients. Half of all trials are delayed, 80 to 90% of them because of a shortage of trial participants. Patient limitations also cause large and unpredicted expenses to pharmaceutical and biotech companies as they are forced to tread water. As the industry moves toward biologics and personalized medicine, this limitation will become even greater. A breakthrough in regulation is needed to create a system that does more with fewer patients.

The current clinical trial system in the United States is more than 50 years old. Its architecture was conceived when electronic manipulation of data was limited, slow, and expensive. Since then, network and connectivity costs have declined ten thousand–fold, data storage costs over a million-fold, and computation costs by an even larger factor. Today, complex and powerful applications like electronic commerce are deployed on a large scale. Amazon.com is a good example. A large database of customers and products form the kernel of its operation. A customer's characteristics (like buying history and preferences) are observed and stored. Customers can be grouped and the buying behavior of any individual or group can be compared with corresponding behavior of others. Amazon track how a group or an individual responds to an outsid (such as advertising).

Once safety is proven,

Once safety is proven, patients could access the medicine in question through qualified physicians.

under their purview. Once safety is proven, patients could access the medicine in question through qualified physicians. Patients' responses to a drug would be stored in a database, along with their medical histories. Patient identity would be protected by biometric identifiers, and the database would be open to qualified medical researchers as a "commons." The response of any patient or group of patients to a drug or treatment would be tracked and compared to those of others in the database who were treated in a different manner or not at all. These comparisons would provide insights into the factors that determine real-life efficacy: how individuals or subgroups respond to the drug. This would liberate drugs from the tyranny of the averages that characterize trial information today. The technology would facilitate such comparisons at incredible speeds and could quickly highlight negative results. As the patient population in the database grows and time passes, analysis of the data would also provide the information needed to conduct postmarketing studies and comparative effectiveness research.

Today's e-commerce systems started small and took nearly 20 years to develop. Adapting this kind of capability to medical information would be a monumental undertaking. Initiating and overseeing it would be an appropriate task for the professional societies. There are encouraging signs, including a call in 2004 by the American Medical Association for public registries of drugs, as well as a proposal for trials that incorporate feed-forward mechanisms.* Another proposal would allow patients to choose between medicines whose efficacy has been determined in different manners.† There is also a suggestion to use control of pricing to encourage drug developers to move forward in a "progressive" trial design.‡ Ideas, however, are not enough. We need the professions to mobilize and take advantage of this enormous opportunity.

– Andrew Grove

10.1126/science.1212118

*H. Hvitfeldt *et al.*, *Qual. Manag. Health Care* **18**, 247 (2009). †B. J. Madden, *Free to Choose Medicine* (Heartland Institute, Chicago, IL, 2010). ‡M. Boldrin, S. J. Swamidass, "A New Bargain for Drug Approvals," *Wall Street Journal*, 25 July 2011.

Another issue of Science Sep23/11

Grove--- former CEO Intel

⇒ immense database for medical researchers

But

Once safety is proven,

isn't so simple!

"Statistics"

Vioxx

- An analgesic (pain relief) from Merck "Pharma"

1999 Approved by FDA

2000 Advertizing $160m .... More than Pepsi or Budweiser

2003 Earned $2.46 b Big bucks!

Vioxx

- An analgesic (pain relief) from Merck "Pharma"

1999 Approved by FDA

2000 Advertizing $160m .... More than Pepsi or Budweiser

2003 Earned $2.46b

- Evidence available: of 3x, 5x rate of CV events (cardio-vascular)

Vioxx

- An analgesic (pain relief) from Merck "Pharma"

  1999 Approved by FDA

  2000 Advertizing $160m .... More than Pepsi or Budweiser

  2003 Earned $2.46b

- Evidence available of 3x, 5x rate of CV events (cardio-vascular)

- Government litigation vs. Merck: 2004

  - that M knew of increased incidence of CV events
  - that M had "filtered" Data
  - that V linked To 30,000 to 40,000 deaths
  - David Madigan, Chair, Columbia U Statistics

Vioxx

- An analgesic (pain relief) from Merck "Pharma"

  1999 Approved by FDA

  2000 Advertizing $160m .... More than Pepsi or Budweiser

  2003 Earned $2.46b

- Evidence available of 3x, 5x rate of CV events (cardio-vascular)

- Government litigation vs. Merck

  - that M knew of increased incidence of CV events
  - that M had "filtered" Data
  - that V linked To 30,000 to 40,000 deaths
  - David Madigan, Chair, Columbia U Statistics

- Merck voluntarily withdrew Vioxx

  - Got $5b penalty
    - A "bargain" as they'd made $10b
  - and FDA estimated 27,787 deaths

Vioxx

- An analgesic (pain relief) from Merck "Pharma"

  1999 Approved by FDA

  2000 Advertizing $160m .... More than Pepsi or Budweiser

  2003 Earned $2.46b

- Evidence available of 3x, 5x rate of CV events (cardio-vascular)

- Government litigation vs. Merck

  - that M knew of increased incidence of CV events
  - that M had "filtered" Data
  - that V linked To 30,000 to 40,000 deaths
  - David Madigan, Chair, Columbia U Statistics

- Merck voluntarily withdrew Vioxx

  - Got $5b penalty
    - A "bargain" as they'd made $10b
  - but FDA estimated 27,787 deaths

- Big $s Big deaths Big Data-obfuscation Big Statistics

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

Replication
again, again, again

6 articles
9 pages

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

Replication
again, again, again

6 articles
9 pages

It was a core of statistics
- once

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

Replication
again, again, again

6 articles
9 pages

Unfortunately, it can also indicate fraud

of statistics
-once upon a time !

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

Replication
again, again, again

6 articles
9 pages

Unfortunately, it can also indicate fraud

-once upon a time !

Maybe!
"Fortunately
it can also indicate fraud"

INTRODUCTION

Again, and Again, and Again ...

REPLICATION—THE CONFIRMATION OF RESULTS AND CONCLUSIONS FROM ONE STUDY obtained independently in another—is considered the scientific gold standard. New tools and technologies, massive amounts of data, long-term studies, interdisciplinary approaches, and the complexity of the questions being asked are complicating replication efforts, as are increased pressures on scientists to advance their research. The five Perspectives in this section (and associated News and Careers stories, Readers' Poll, and Editorial) explore some of the issues associated with replicating results across various fields.

Ryan (p. 1229) highlights the excitement and challenges that come with field-based research. In particular, observing processes as they occur in nature allows for discovery but makes replication difficult, because the precise conditions surrounding the observations are unique. Further, although laboratory research allows for the specification of experimental conditions, the conclusions may not apply to the real world. Debate about the merits of lab-based and field-based studies has been a persistent theme over time. Tomasello and Call (p. 1227) further contribute to this debate in their discussion of some obvious barriers to replication in primate cognition and behavior research (small numbers of subjects, expense, and ethics issues) as well as more subtle ones, such as the nontrivial challenge of designing tasks that elicit complex cognitive behaviors.

New technologies continue to produce a deluge of data of different varieties, raising expectations for new knowledge that will translate into meaningful therapeutics and insights into health. Ioannidis and Khoury (p. 1230) outline multiple steps for validating such large-scale data on the path to clinical utility and make suggestions for incentives (and penalties) that could enhance the availability of reliable data and analyses.

Peng (p. 1226) discusses the need for a minimum standard of reproducibility in computer sciences, arguing that enough information about methods and code should be available for independent researchers to reach consistent conclusions using original raw data. Specifically, he describes a model that one journal has used to make this a reality.

The need to convince the public that data are replicable has grown as science and public policy-making intersect, an issue that has has beset climate change studies. As Santer *et al.* (p. 1232) describe, having multiple groups examining the same data and generating new data has led to robust conclusions.

The importance of replication and reproducibility for scientists is unquestioned. Sometimes attempts to replicate reveal scientific uncertainties. This is one of the main ways that science progresses (see associated News stories of faster-than-light neutrinos and sirtuins, pp. 1200 and 1194). Unfortunately, in rare instances (compared to the body of scientific work), it can also indicate fraud (see the Editorial by Crocker and Cooper, p. 1182). How do we promote the publication of replicable data? The authors in this section come up with possibilities that are targeted at funders, journals, and the research culture itself. In the Readers' Poll, you can make your views known as well.

– BARBARA R. JASNY, GILBERT CHIN, LISA CHONG, SACHA VIGNIERI

Data Replication & Reproducibility

CONTENTS

Perspectives

1226 Reproducible Research in Computational Science
R. D. Peng

1227 Methodological Challenges in the Study of Primate Cognition
M. Tomasello and J. Call

1229 Replication in Field Biology: The Case of the Frog-Eating Bat
M. J. Ryan

1230 Improving Validation Practices in "Omics" Research
J. P. A. Ioannidis and M. J. Khoury

1232 The Reproducibility of Observational Estimates of Surface and Atmospheric Temperature Change
B. D. Santer et al.

See also Editorial p. 1182; News stories pp. 1194 and 1200; Readers' Poll p. 1203; Science Careers content p. 1179; and www.sciencemag.org/special/data-rep/

Science

Science again

Replication
again, again, again

6 articles
9 pages

Unfortunately, of statistics
it can also indicate fraud
-once upon a time !

Maybe!
Fortunately
it can also indicate fraud

If fraud is there,
we ought to know about it!

Messages ?

Data : - No role for statistics

Messages ?

Data : - No role for statistics

Groves : Drug deemed Safe - Available by prescriptions
- Collect data ... massively

Messages ?

Data : - No role for statistics

Groves : Drug deemed Safe
- Available by prescriptions
- Collect data ... massively

Vioxx Drug deemed Safe
- 27,787 deaths
- $5b profit out of $10b

Messages ?

Data : - No role for statistics

Groves: Drug deemed Safe
- Available by prescriptions
- Collect data ... massively

Vioxx: Drug deemed Safe
- 27,787 deaths
- $5b profit out of $10b

Replication: Double check those claims !

Messages ?

Data : - No role for statistics

Groves : Drug deemed Safe - Available by prescriptions
- Collect data ... massively

Vioxx Drug deemed Safe - 27,787 deaths
- $5b profit out of $10b

Replication : Double check those claims !

Conflicting messages BIG scale

Overt statistics

Hidden statistics

Business as usual

The principles ?

Messages ?

Data : - No role for statistics

Groves : Drug deemed Safe
- Available by prescriptions
- Collect data ... massively

Vioxx Drug deemed Safe
- 27,787 deaths
- $5b profit out of $10b

Replication : Double check those claims !

Conflicting messages BIG scale

Overt statistics

Hidden statistics

Business as usual

The principles ?

Is the issue:

public relations ?

or the

discipline itself ?

(b) Statistics: Two logics

How can that happen? How to answer?

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn:

B Invert density fn:

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

B Invert density fn:

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

B Invert density fn: need to get prior $\pi(\theta)$

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

documentation (it's there)

B Invert density fn: need to get prior $\pi(\theta)$

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

documentation (it's there)

B Invert density fn: need to get prior $\pi(\theta)$

various versions

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

documentation (it's there)

B Invert density fn: need to get prior $\pi(\theta)$

various versions

Two logics ? Various versions ?

(b) Statistics: Two logics

How can that happen? How to answer?

A Invert dist'n fn: need to get pivot

documentation (it's there)

B Invert density fn: need to get prior $\pi(\theta)$

various versions

Two logics ? Various versions ?

As a discipline:

Can we afford it? "PR" or "substance"

Who else has two logics ?

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ — Been around as long as prob. theory

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: - Combine the prob's
- Record separately

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: - Combine the prob's
- Record separately

Nothing says you have to combine!

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: - Combine the prob's
- Record separately

Nothing says you have to combine!

(ii) $\pi(\theta)$ introduced for convenience

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: Combine the probs
Record separately

Nothing says you have to combine!

(ii) $\pi(\theta)$ introduced
for convenience

Where Bayes started in 1763!

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: - Combine the prob's
- Record separately

Nothing says you have to combine!

(ii) $\pi(\theta)$ introduced for convenience

where Bayes started in 1763!

and (i) was "co-opted" for "turf" for ---

(c) What are the versions?

(i) $\pi(\theta)$ describes source of θ

Been around as long as prob. theory

Options: - Combine the prob's
- Record separately

Nothing says you have to combine!

(ii) $\pi(\theta)$ introduced for convenience

Where Bayes started in 1763!

and (i) was "co-opted" for "turf" for ---

Again Can we afford it? "PR" or "substance"
Who else has two logics? Versions?

Versions:

(ii) for <u>convenience</u> : the real Bayes ...

Bayes (1763)	original
Laplace (1820)	onto something good
Jeffreys (1946)	mechanize : use $\pi = \lvert i(\theta) \rvert^{1/2}$ (root information)
Jeffreys (1961)	modify
Bernardo (1979)	reference priors (measure spaces)
... more recent ...	(later)

Can the right $\pi(\theta)$ do it all ?

(d) Can the "right" $\pi(\theta)$ do it all?

Need a <u>vector</u> parameter $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

(d) Can the "right" $\pi(\theta)$ do it all?

Need a vector parameter $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

Try: Normal on the plane

$y_1 = \theta_1 + z_1$

$y_2 = \theta_2 + z_n$

z_i: N o i

(d) Can the "right" $\pi(\theta)$ do it all?

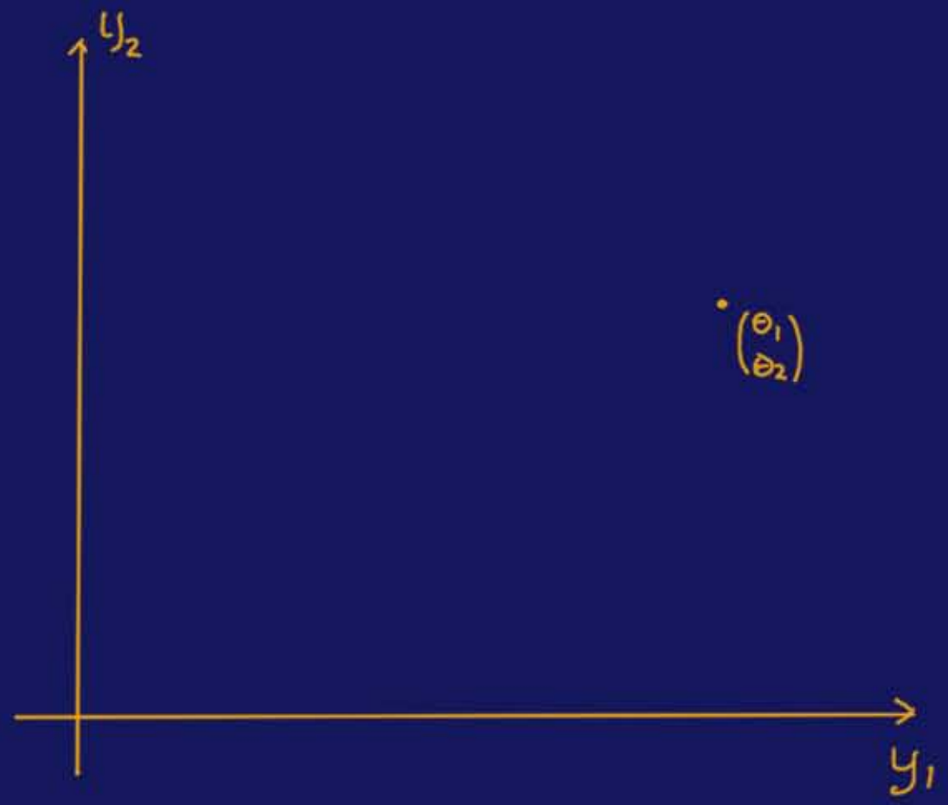
Need a vector parameter $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

Try: Normal on the plane

$y_1 = \theta_1 + z_1$

$y_2 = \theta_2 + z_n$

z_i: N o i



(d) Can the "right" $\pi(\theta)$ do it all?

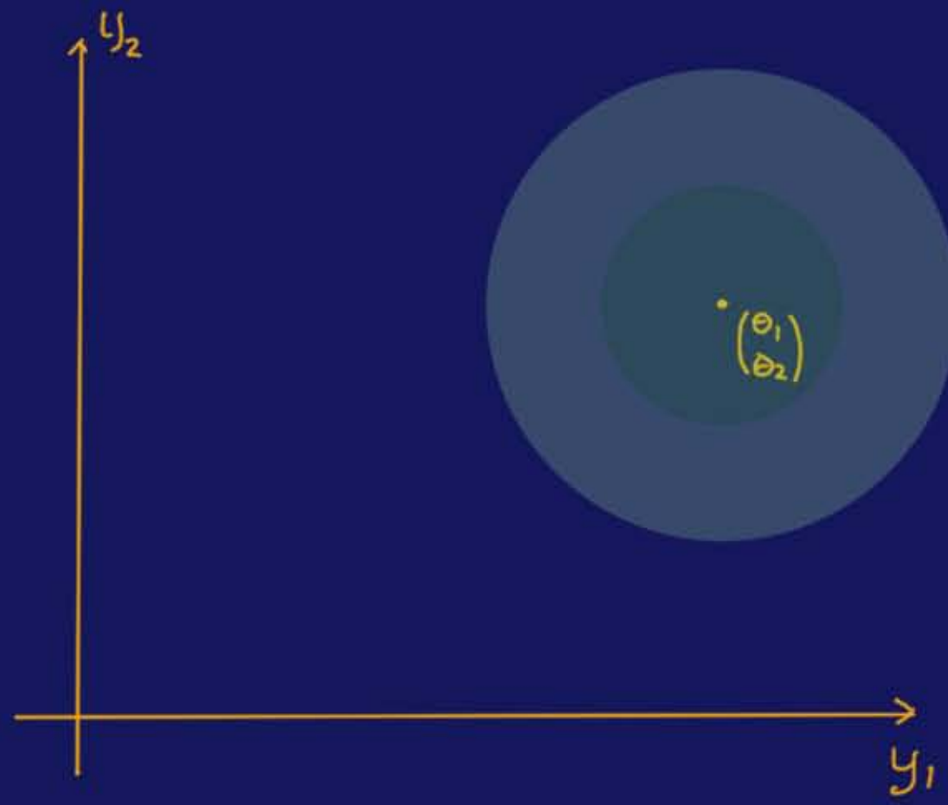
Need a vector parameter $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

Try: Normal on the plane

$y_1 = \theta_1 + z_1$

$y_2 = \theta_2 + z_n$

z_i: N o i



(d) Can the "right" $\pi(\theta)$ do it all?

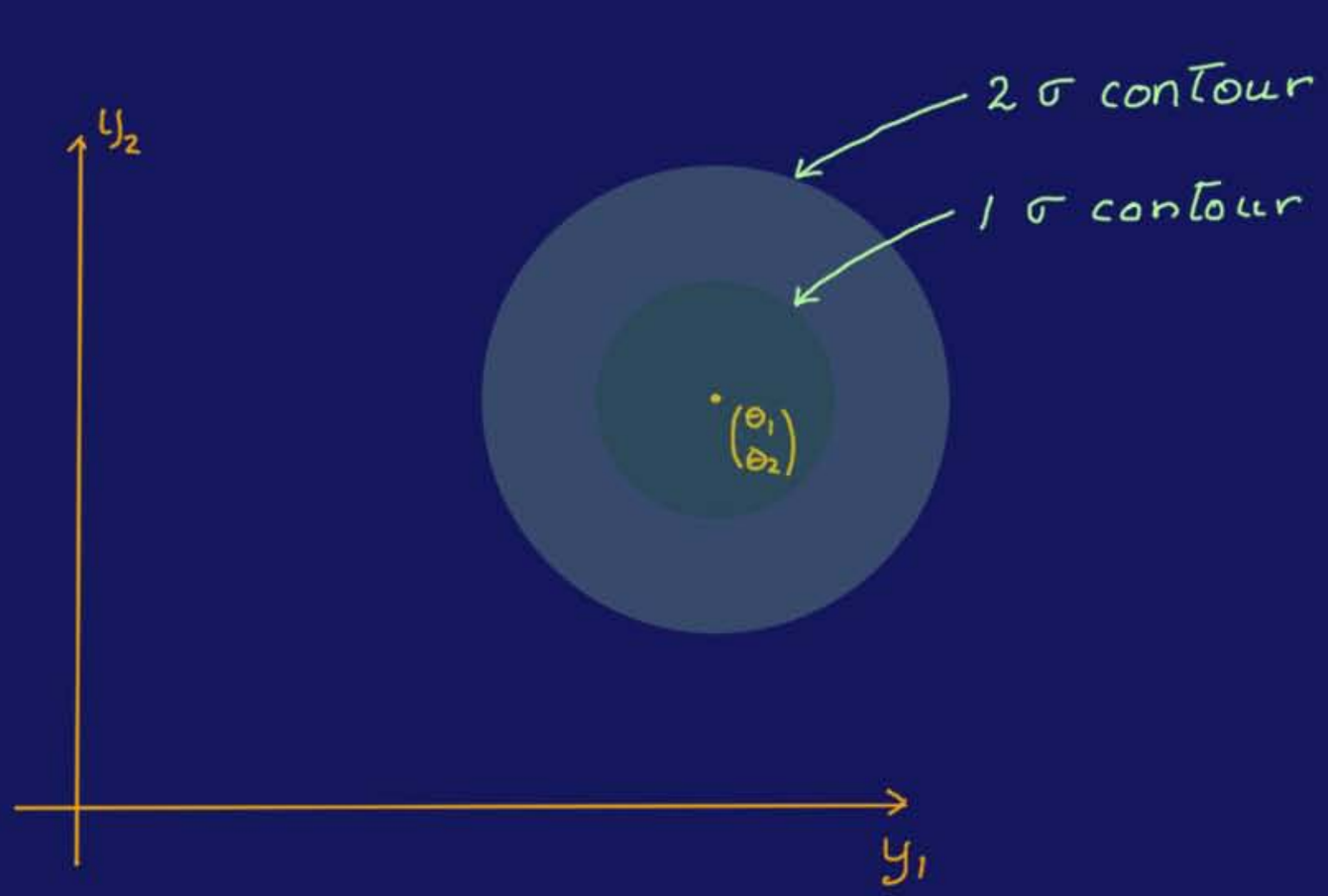
Need a vector parameter $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

Try: Normal on the plane

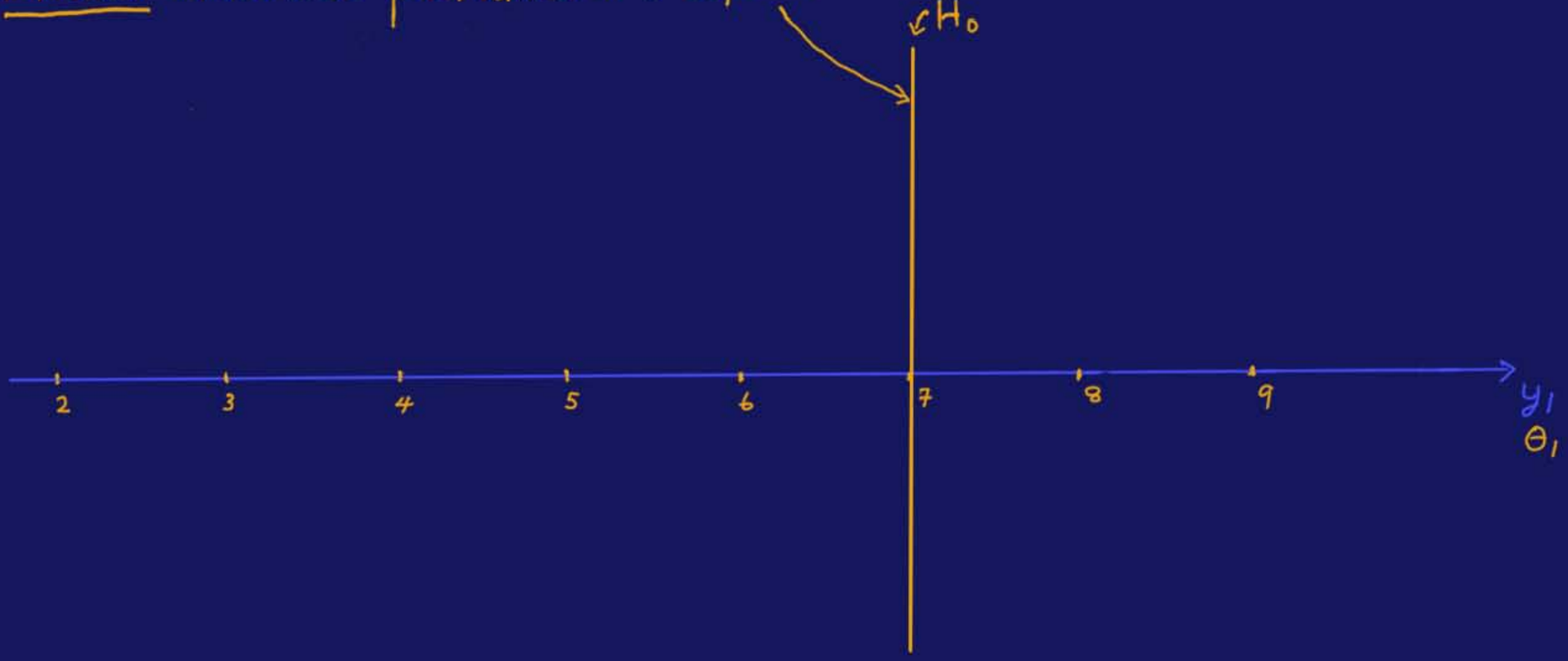
$y_1 = \theta_1 + z_1$

$y_2 = \theta_2 + z_n$

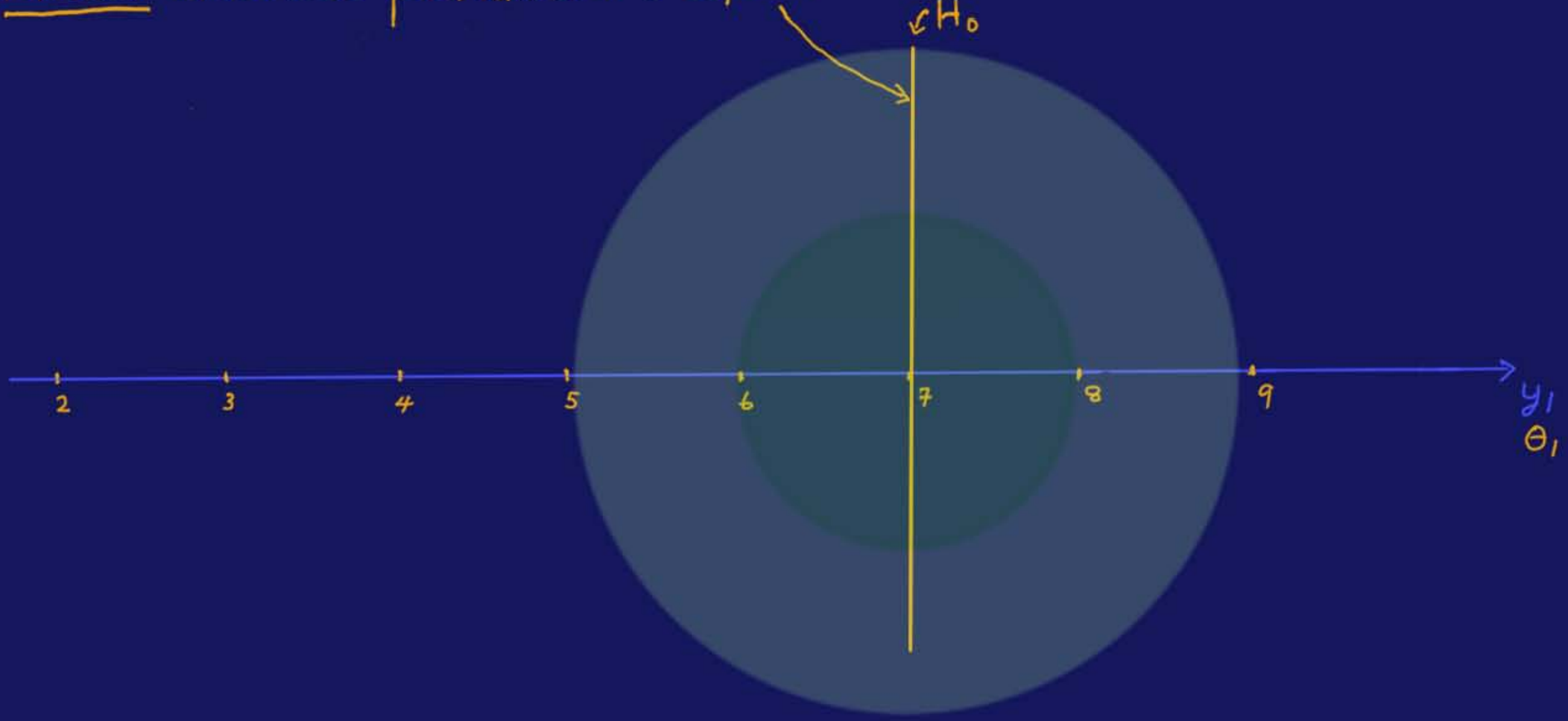
z_i: N o i



(i) Linear interest parameter: $\theta_1 = 7$

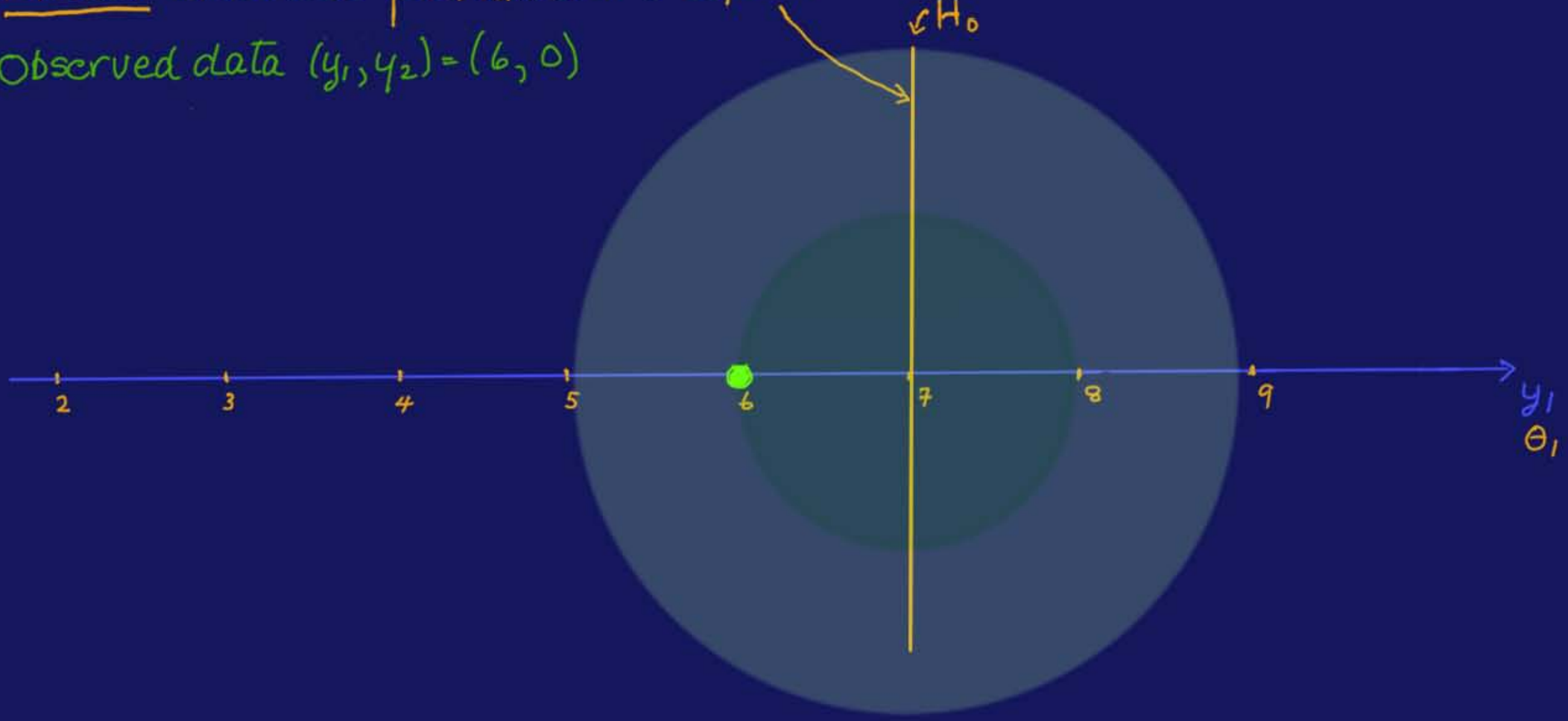


(i) Linear interest parameter: $\theta_1 = 7$



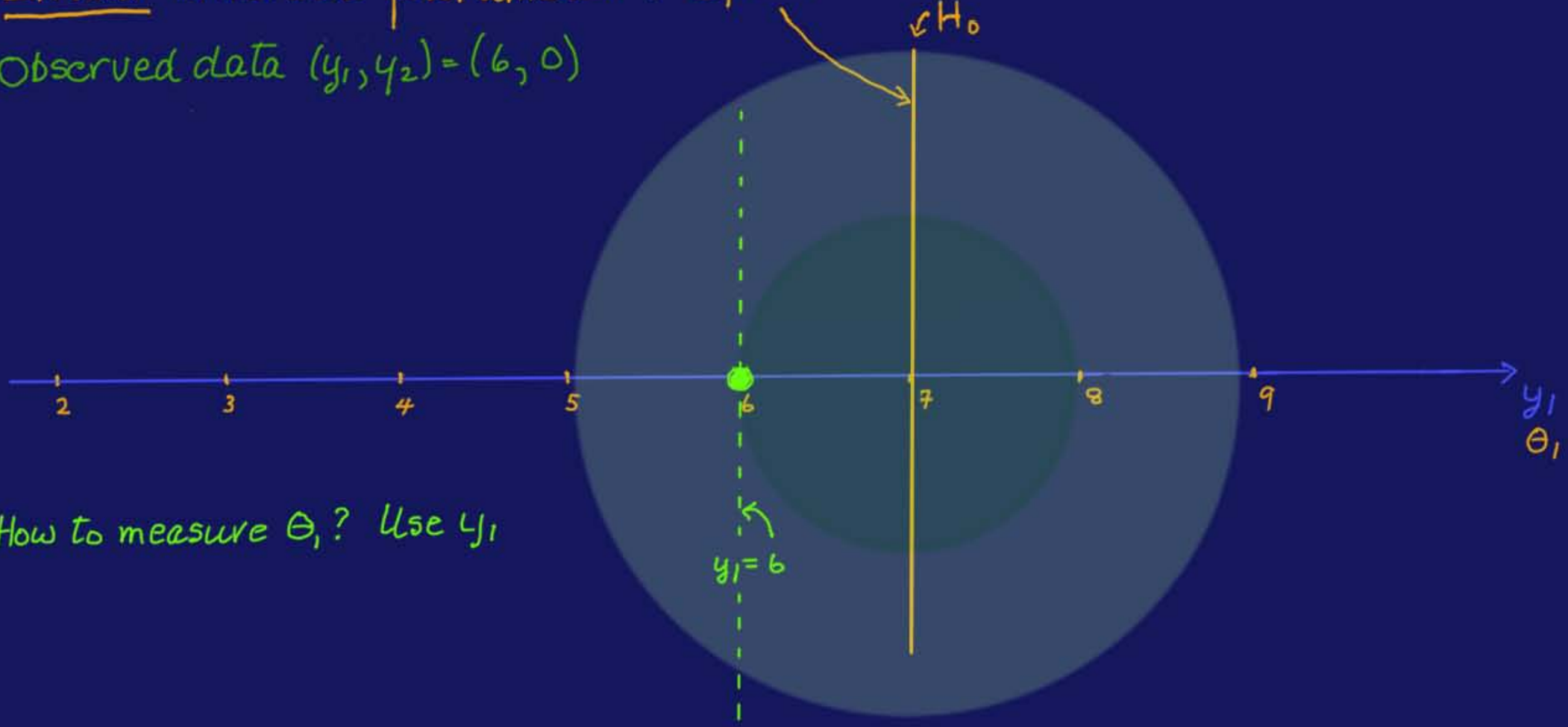
(i) Linear interest parameter: $\theta_1 = 7$

Observed data $(y_1, y_2) = (6, 0)$



(i) Linear interest parameter: $\theta_1 = 7$

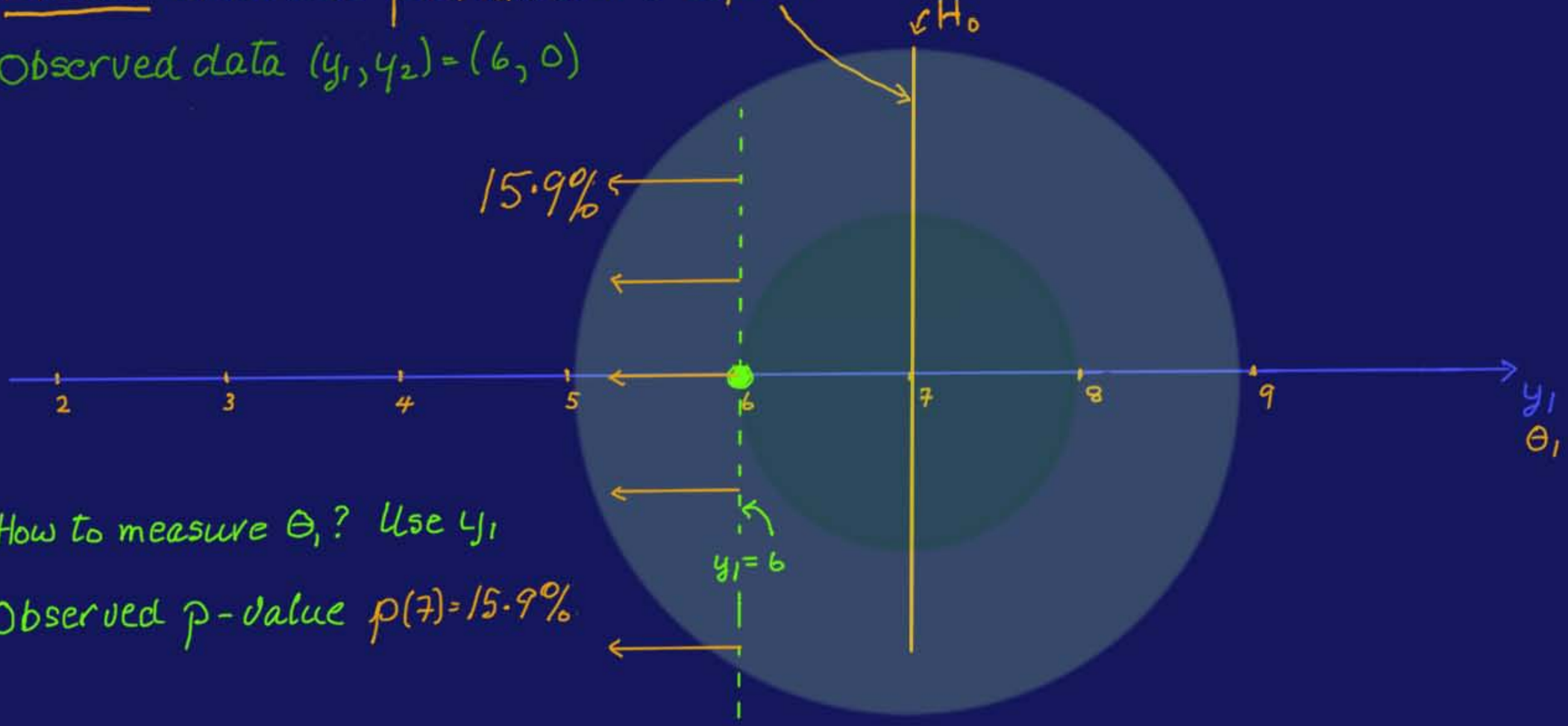
Observed data $(y_1, y_2) = (6, 0)$



How to measure θ_1? Use y_1

(i) Linear interest parameter: $\theta_1 = 7$

Observed data $(y_1, y_2) = (6, 0)$

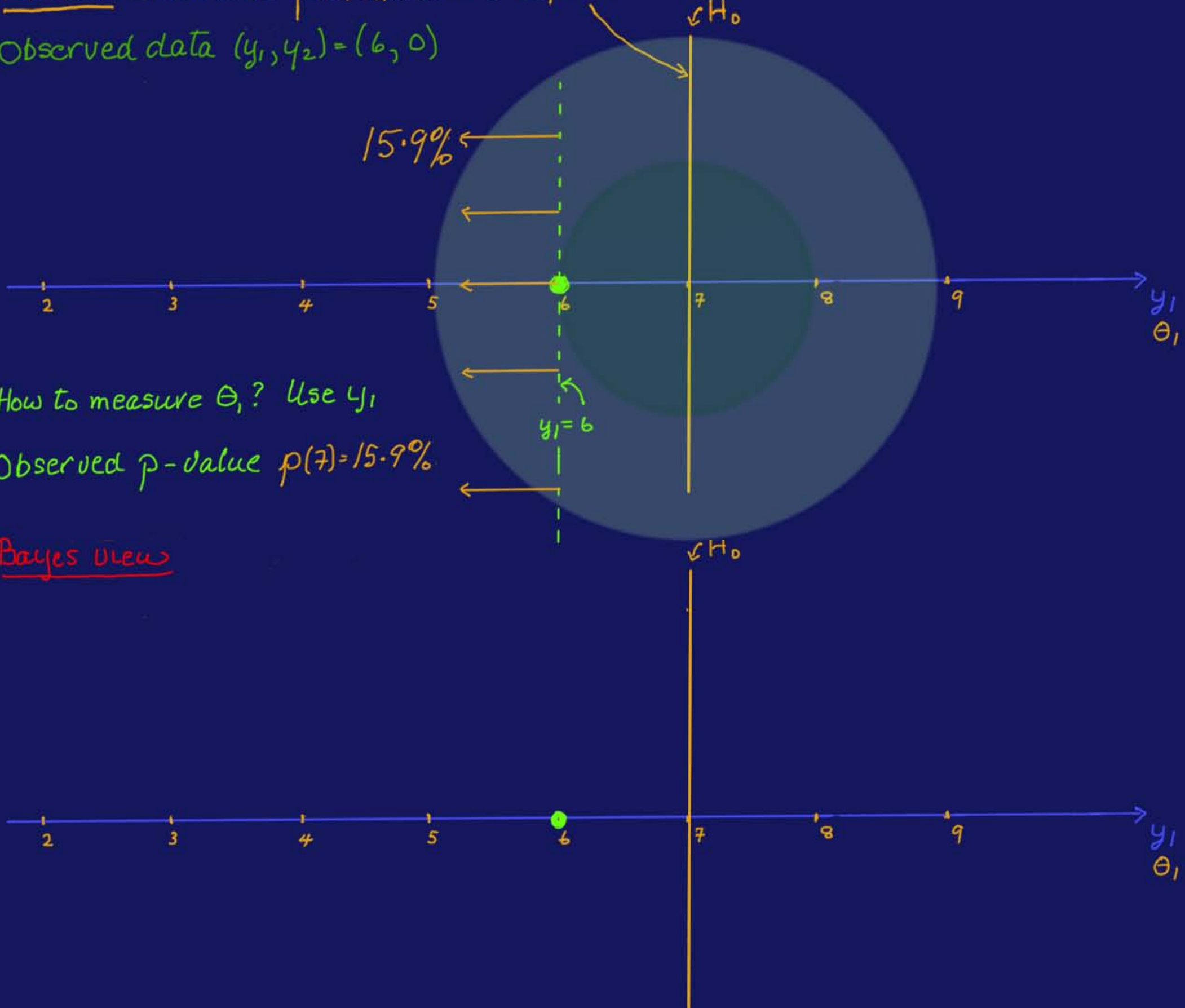


How to measure θ_1? Use y_1

Observed p-value $p(7) = 15.9\%$

(i) Linear interest parameter: $\theta_1 = 7$

Observed data $(y_1, y_2) = (6, 0)$



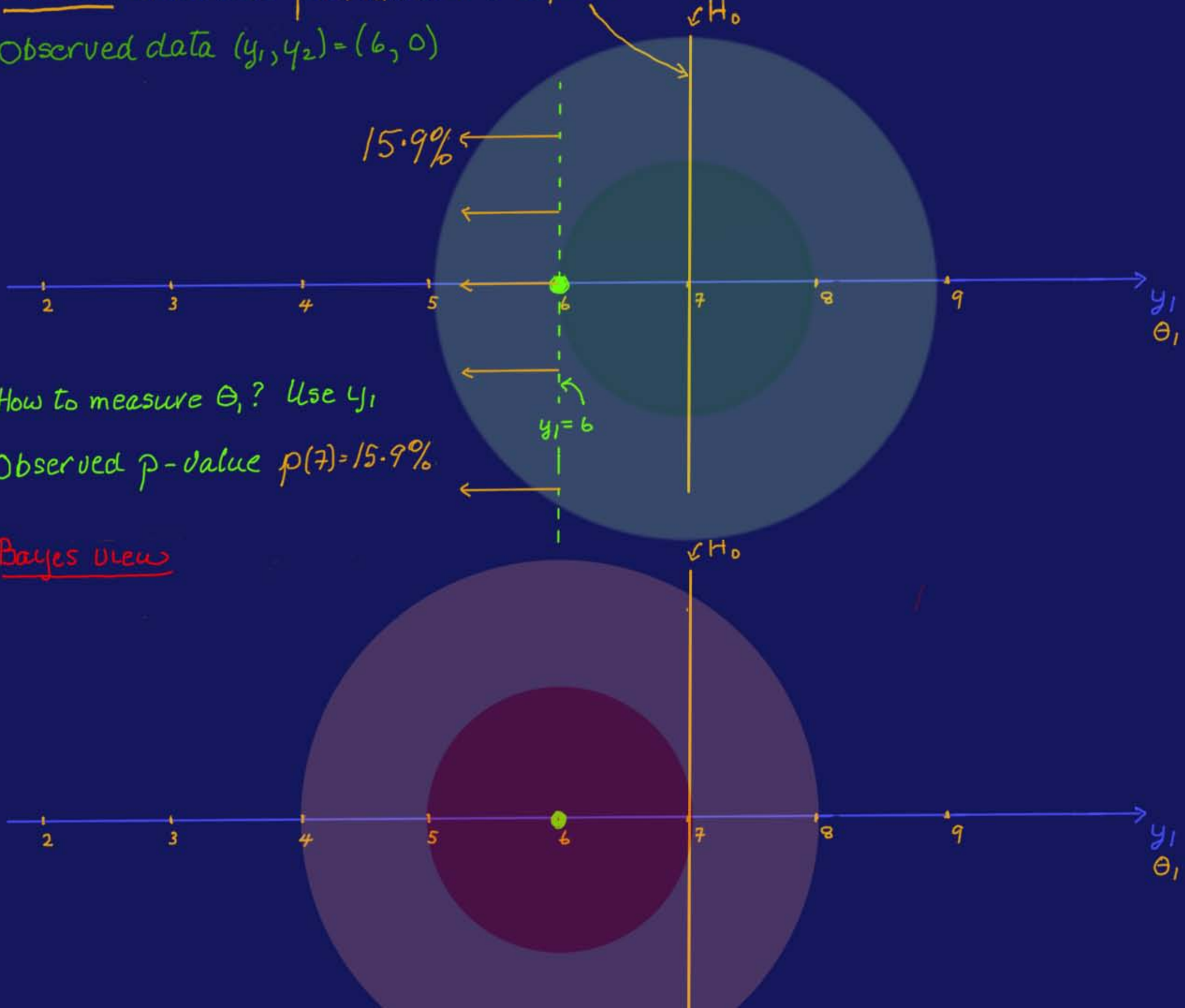
How to measure θ_1? Use y_1

Observed p-value $p(7) = 15.9\%$

Bayes view

(i) Linear interest parameter: $\theta_1 = 7$

Observed data $(y_1, y_2) = (6, 0)$



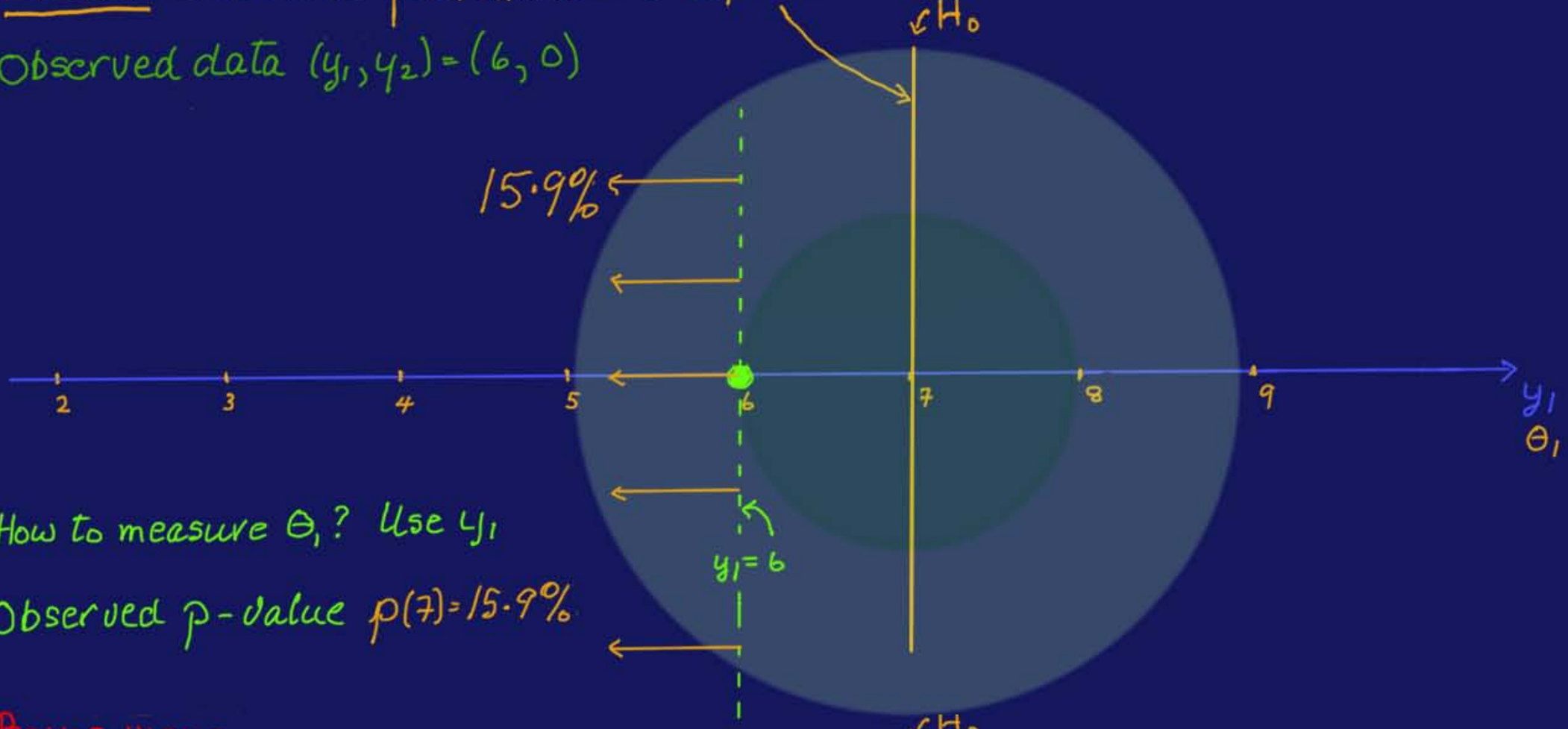
How to measure θ_1? Use y_1

Observed p-value $p(7) = 15.9\%$

Bayes view

(i) Linear interest parameter: $\theta_1 = 7$

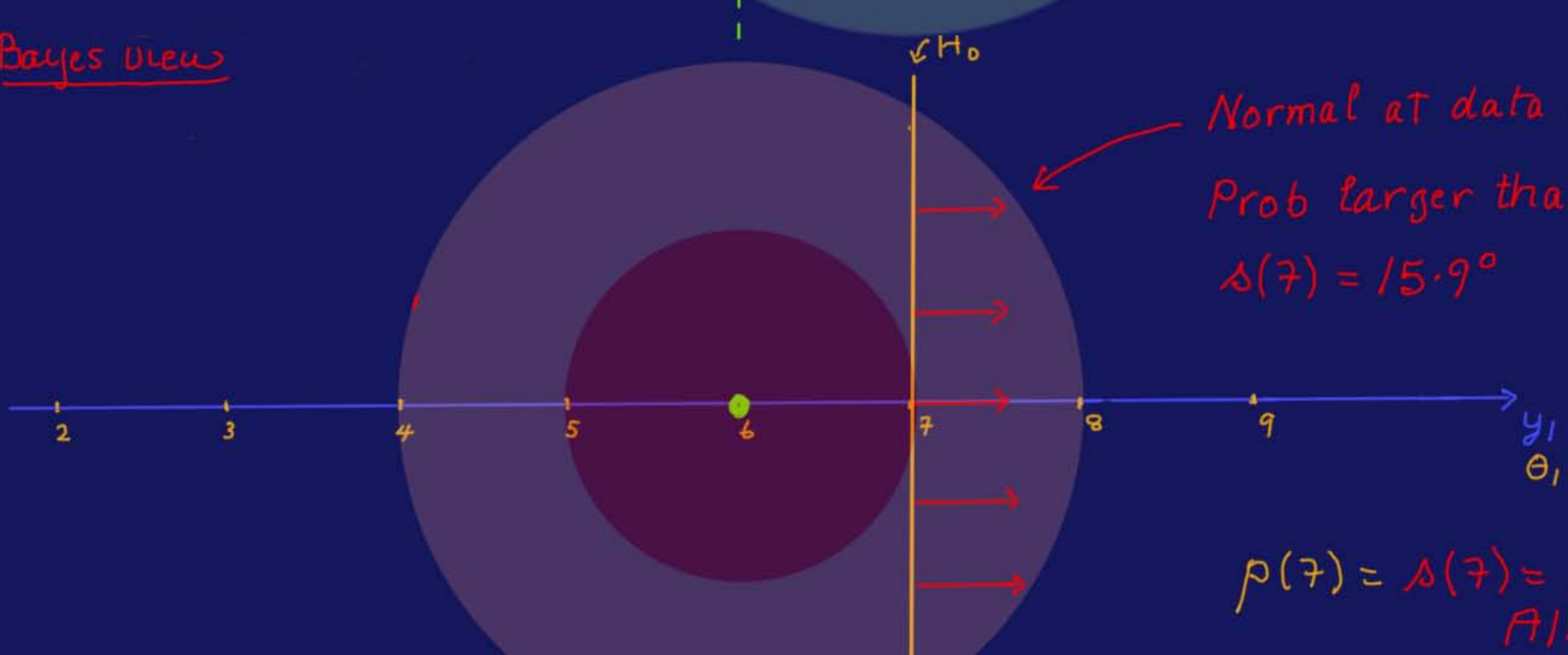
Observed data $(y_1, y_2) = (6, 0)$



How to measure θ_1? Use y_1

Observed p-value $p(7) = 15.9\%$

Bayes view



Normal at data $(6, 0)$

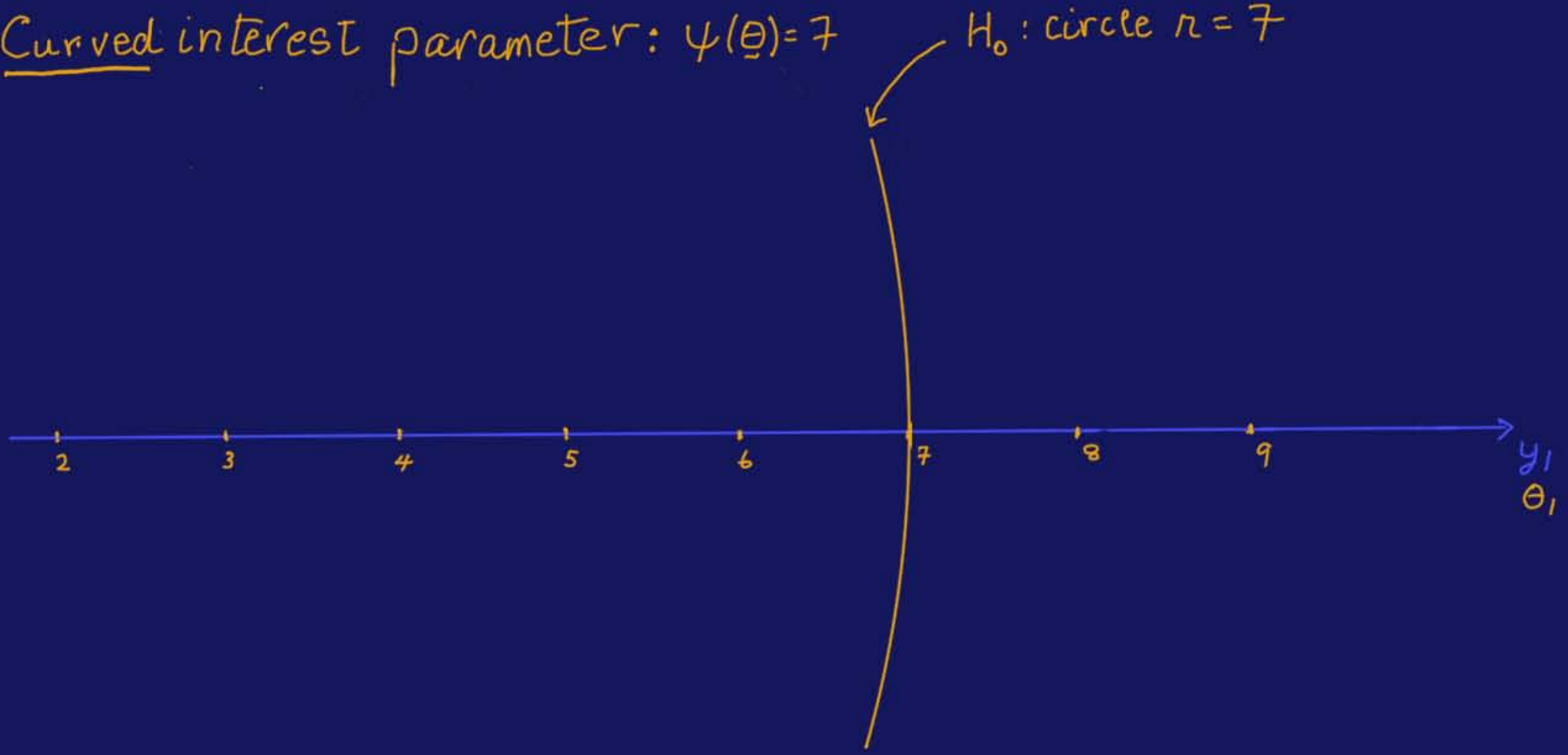
Prob larger than $\theta_1 = 7$

$s(7) = 15.9\%$

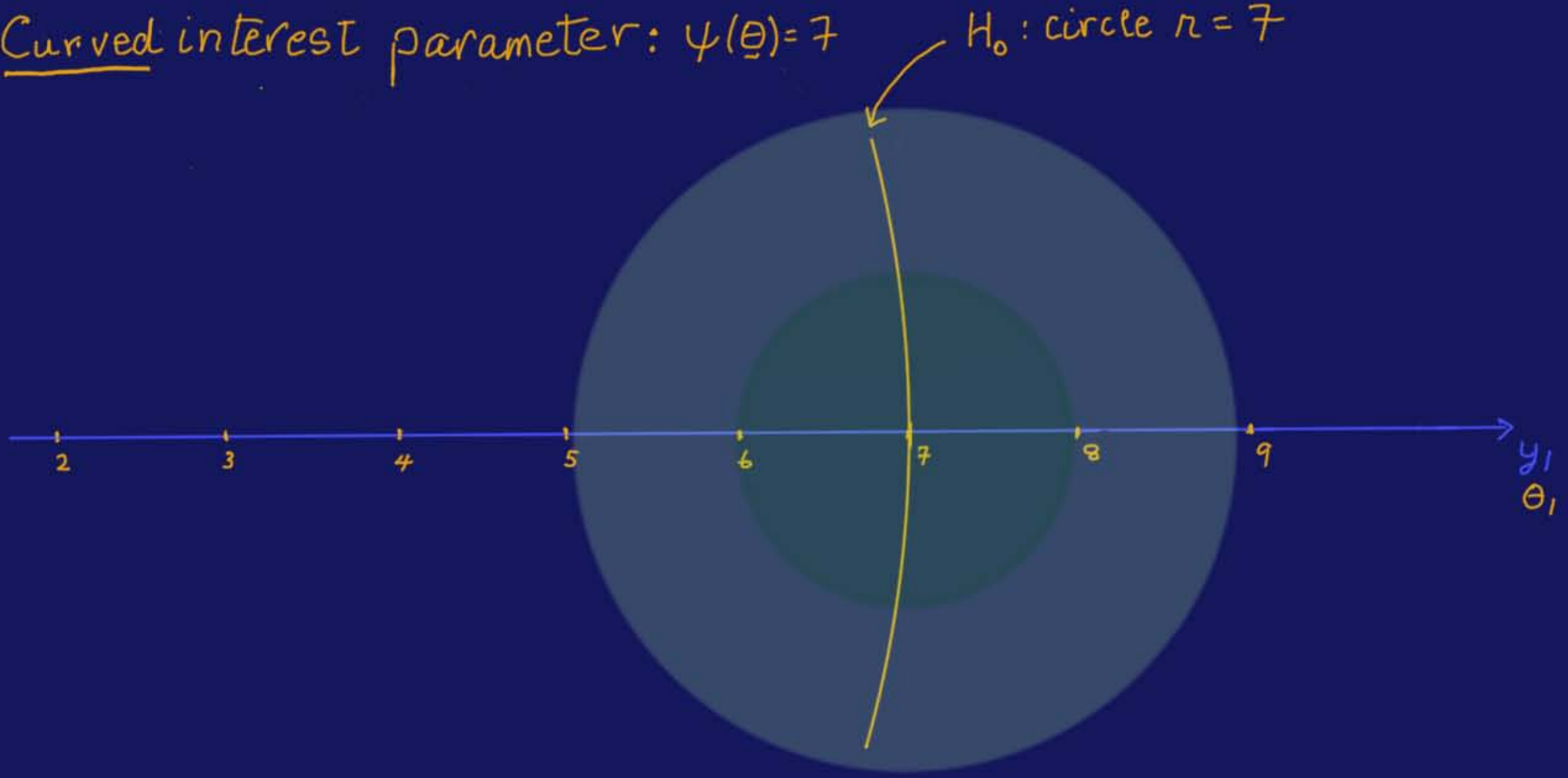
$p(7) = s(7) = 15.9\%$

All OK!

(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$



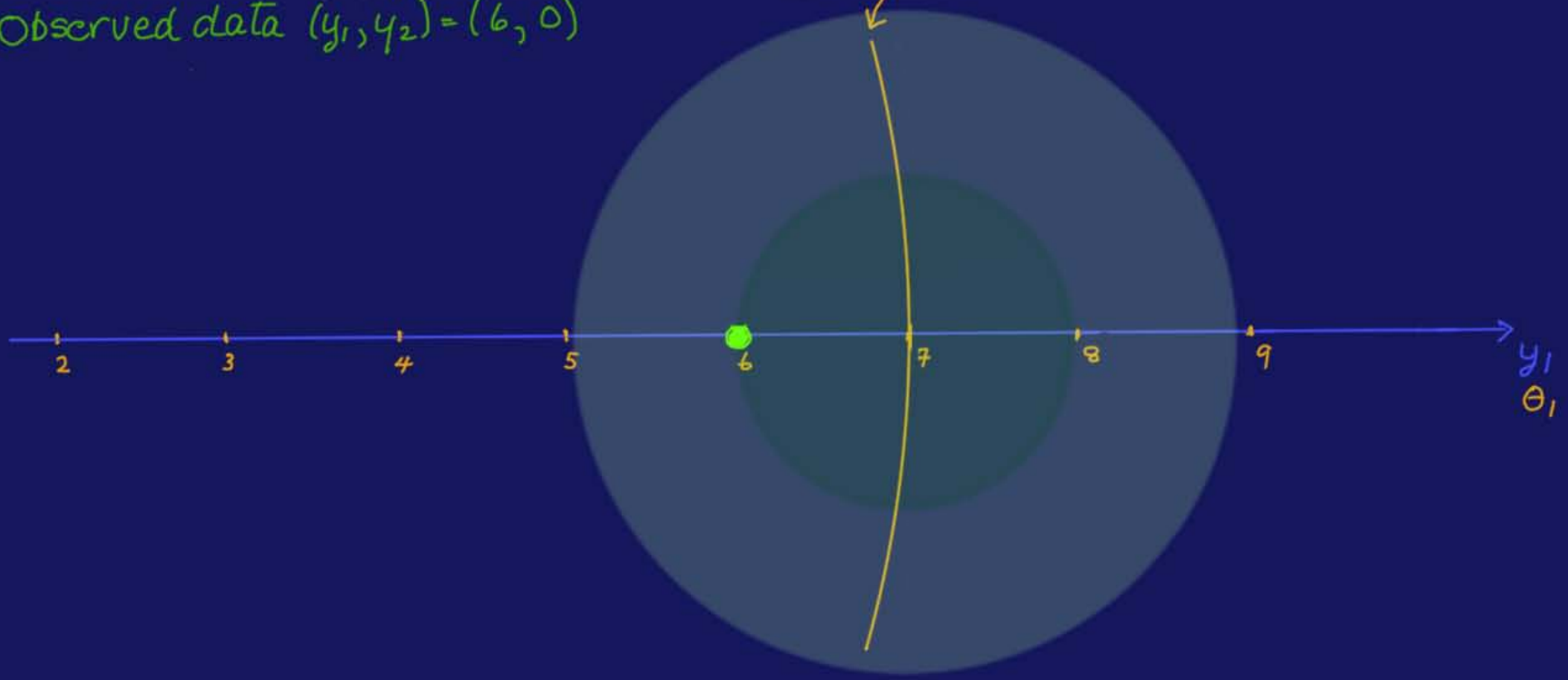
(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$



(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

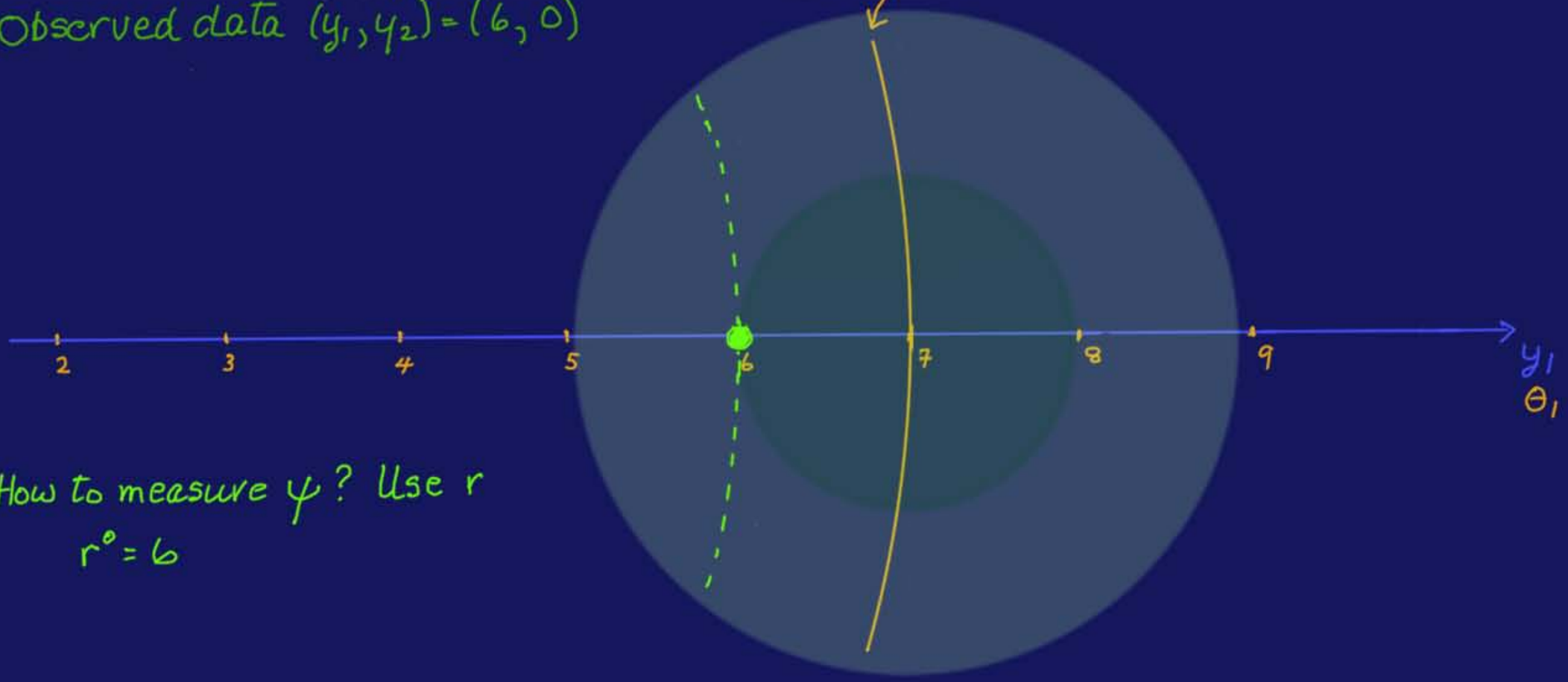
Observed data $(y_1, y_2) = (6, 0)$



(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



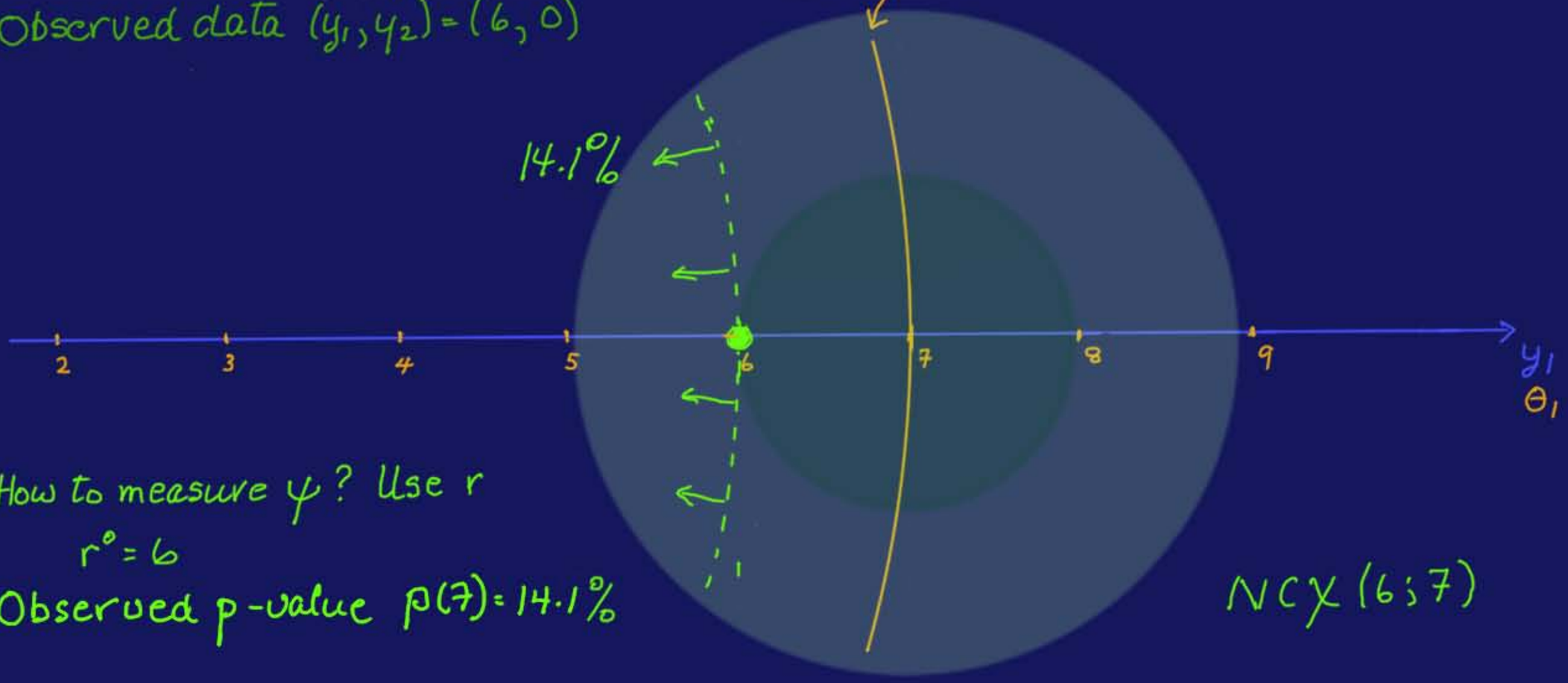
How to measure ψ? Use r

$r^o = 6$

(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



How to measure ψ? Use r

$r^o = 6$

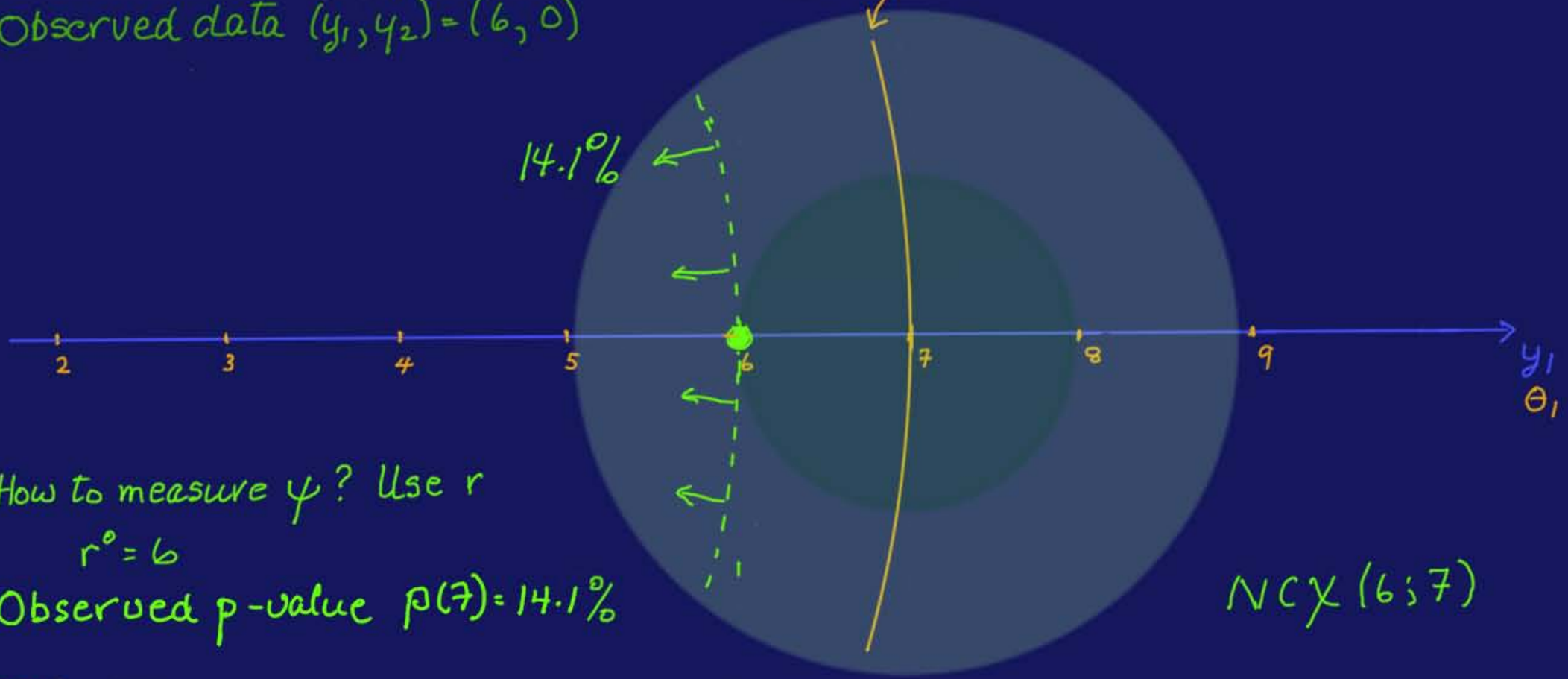
Observed p-value $p(7) = 14.1\%$

NCX (6; 7)

(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



How to measure ψ? Use r

$r^o = 6$

Observed p-value $p(7) = 14.1\%$

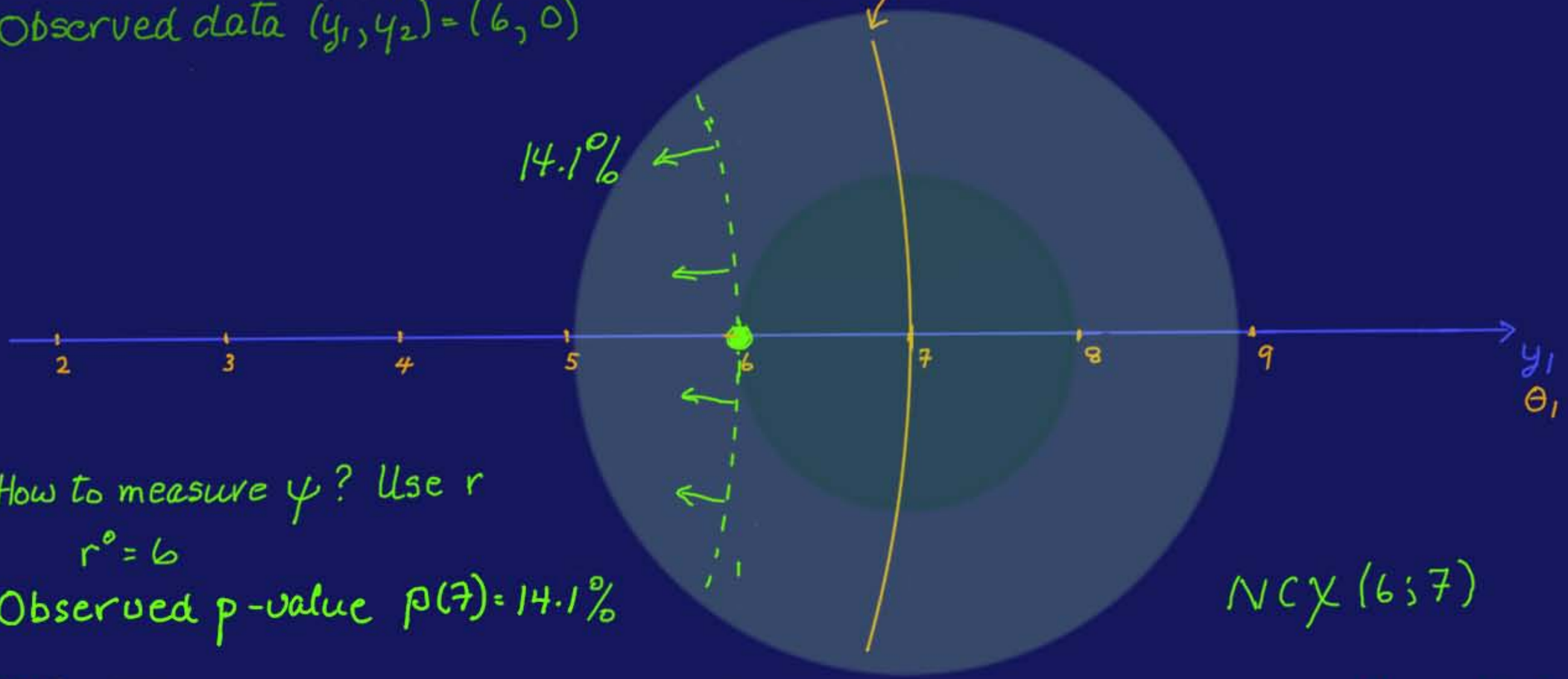
$NCX(6;7)$

Bayes view

(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



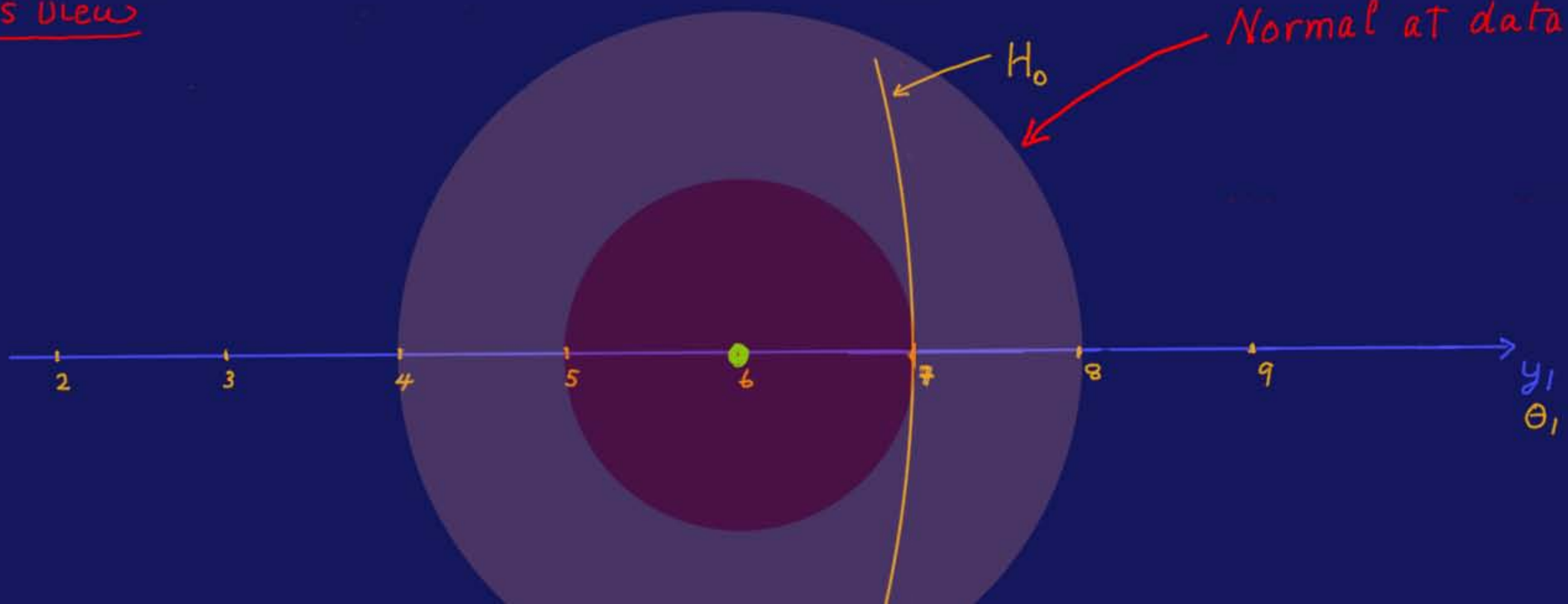
How to measure ψ? Use r

$r^o = 6$

Observed p-value $p(7) = 14.1\%$

$NCX(6;7)$

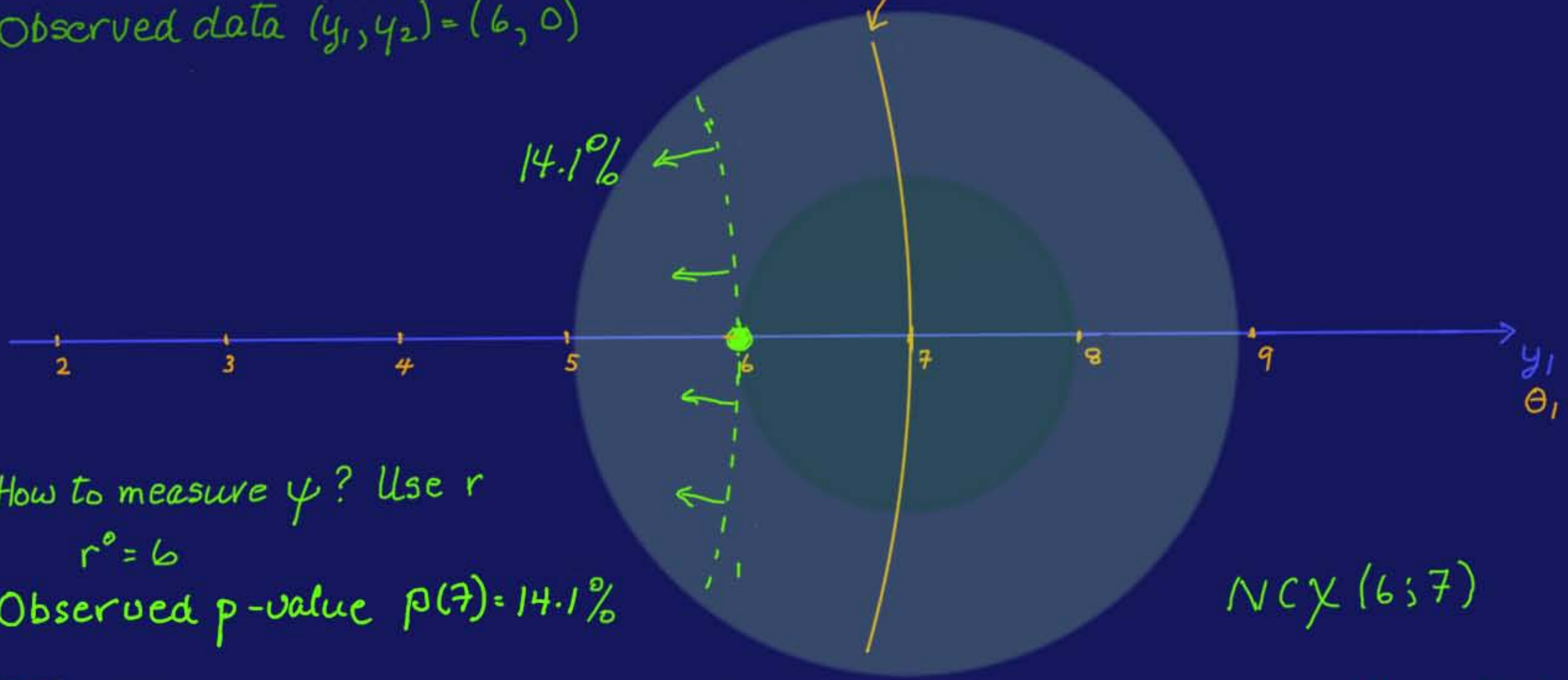
Bayes view



(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



How to measure ψ? Use r

$r^o = 6$

Observed p-value $p(7) = 14.1\%$

NCX (6; 7)

Bayes view

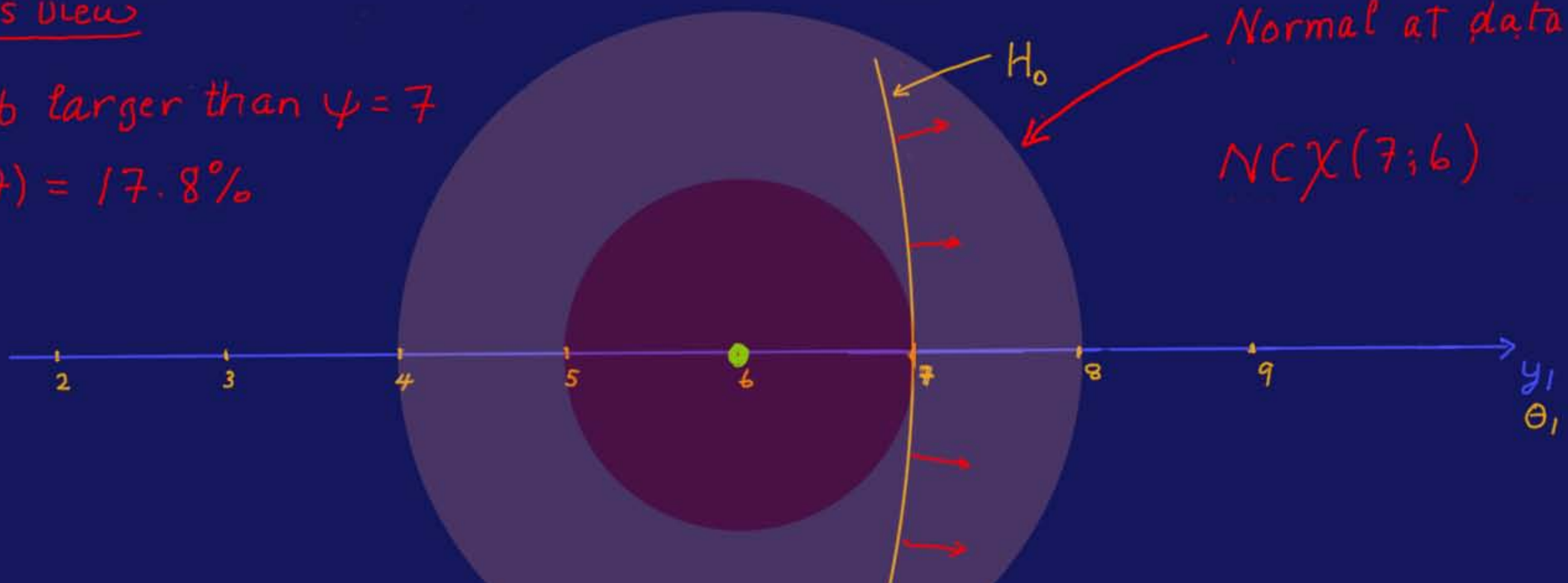
Prob larger than $\psi = 7$

$s(7) = 17.8\%$

Normal at data (6, 0)

H_0

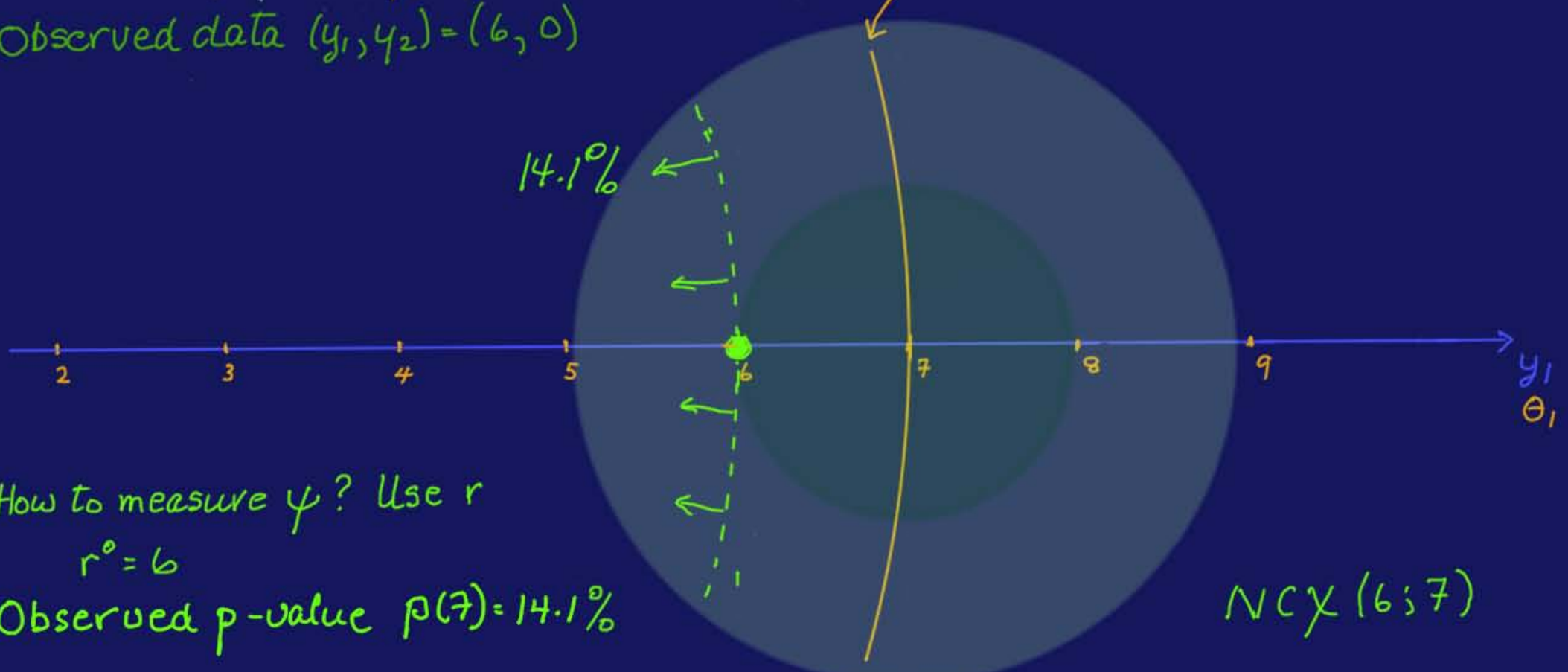
NCX (7; 6)



(e) Curved interest parameter: $\psi(\underline{\theta}) = 7$

H_0 : circle $r = 7$

Observed data $(y_1, y_2) = (6, 0)$



How to measure ψ? Use r

$r^o = 6$

Observed p-value $p(7) = 14.1\%$

NCX (6; 7)

Bayes view

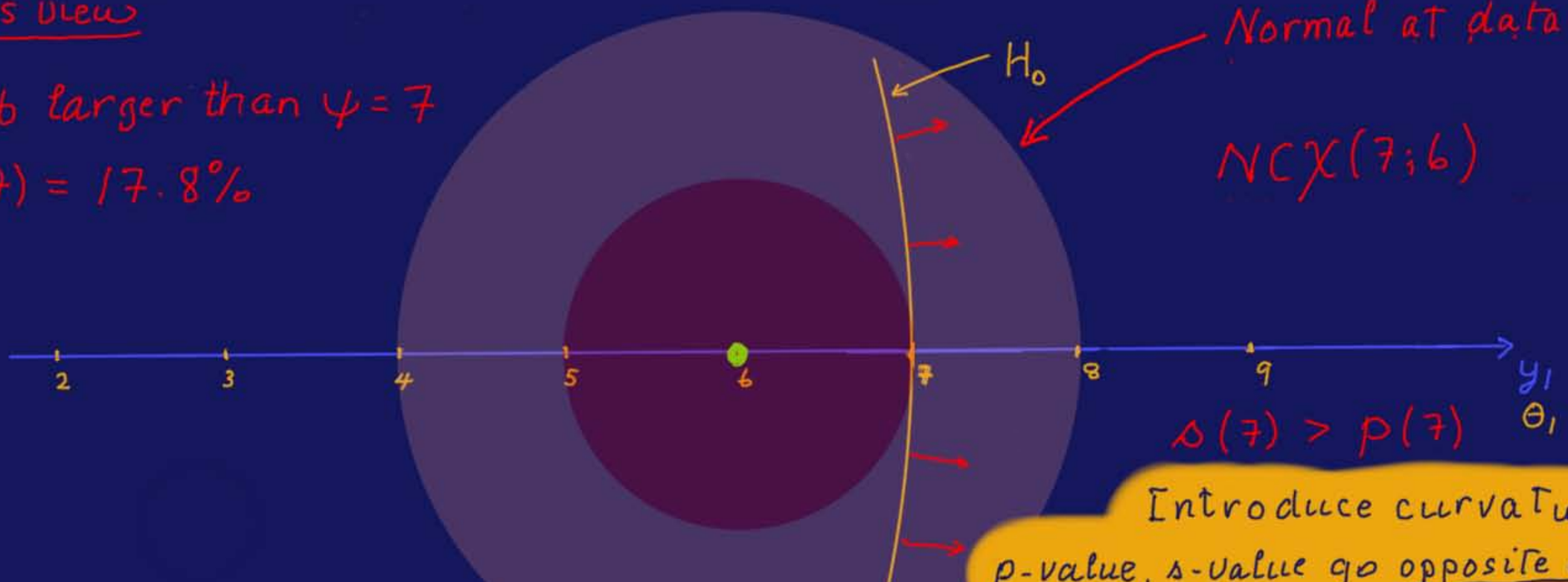
Prob larger than $\psi = 7$

$s(7) = 17.8\%$

Normal at data (6, 0)

H_0

NCX (7; 6)



$s(7) > p(7)$

Introduce curvature:
p-value, s-value go opposite directions

What did that say?

1) Linear parameter of interest

frequency p-value
Bayes s-value
) are equal

What did that say ?

1) Linear parameter of interest

frequency p-value
Bayes s-value
) are equal

2) Introduce parameter curvature (say +)

frequency p-value decreases !
Bayes s-value increases !

What did that say?

1) Linear parameter of interest

frequency p-value
Bayes s-value
} are equal

2) Introduce parameter curvature (say +)

frequency p-value decreases !
Bayes s-value increases !

Which, if either, is right?

p-value has repetition validity !

Confidence distribution
Bayes "
} Can't handle curvature

What did that say?

1) Linear parameter of interest

frequency p-value
Bayes s-value
} are equal

2) Introduce parameter curvature (say +)

frequency p-value decreases !
Bayes s-value increases !

Which, if either, is right?

p-value has repetition validity !

Confidence distribution
Bayes "
} Can't handle curvature

Dawid Stone Zidek 1973 F Reid 2002

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\quad z \sim N(0, I)$

(ii) General: $y = y(\theta, z)$ $\quad z \sim D$

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\qquad z \sim \mathcal{D}(0, I)$

(ii) General: $y = y(\theta, z)$ $\qquad z \sim \mathcal{D}$

Continuity $\Rightarrow$ $\left.\frac{dy}{d\theta}\right|_{y^\circ} = V(\theta)$ = How parameter affects dist'n at y°

$n \times p$

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\quad z \sim \mathcal{D}(0, I)$

(ii) General: $y = y(\theta, z)$ $\quad z \sim \mathcal{D}$

Continuity $\Rightarrow$ $\underset{n \times p}{\left.\frac{dy}{d\theta}\right|_{y^\circ} = V(\theta)}$ = How parameter affects dist'n at y°

Prior: $\pi(\theta)d\theta = |V(\theta)|\,d\theta$

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\qquad z \sim D(0, I)$

(ii) General: $y = y(\theta, z)$ $\qquad z \sim D$

Continuity $\Rightarrow$ $\left.\frac{dy}{d\theta}\right|_{y^\circ} = \underset{n \times p}{V(\theta)}$ = How parameter affects dist'n at y°

Prior: $\pi(\theta)d\theta = |V(\theta)|\,d\theta$

Ex: $y = X\beta + \sigma z$ $\qquad V(\theta) = (X, \dot{z}^\circ(\theta))$ $\qquad \dot{z}^\circ(\theta) = (y^\circ - X\beta)/\sigma$

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\qquad z \sim N(0, I)$

(ii) General: $y = y(\theta, z)$ $\qquad z \sim N$

Continuity $\Rightarrow$ $\underset{n \times p}{\left.\frac{dy}{d\theta}\right|_{y^\circ}} = V(\theta)$ = How parameter affects dist'n at y°

Prior: $\pi(\theta)\,d\theta = |V(\theta)|\,d\theta$

Ex: $y = X\beta + \sigma z$ $\qquad V(\theta) = (X, \overset{\circ}{z}(\theta))$ $\qquad \overset{\circ}{z}(\theta) = (y^\circ - X\beta)/\sigma$

Prior: $\boxed{\pi(\theta)\,d\theta = d\beta\, d\sigma/\sigma}$ $\qquad$ (Jeffreys original $d\beta\, d\sigma/\sigma^p$)

(f) Priors from Continuity

Regular models: "2nd order" theory says prior widely available

Case: Have a quantile representation

(i) Regression: $y = X\beta + \sigma z$ $\qquad z \sim N(0, I)$

(ii) General: $y = y(\theta, z)$ $\qquad z \sim D$

Continuity $\Rightarrow$ $\underset{n\times p}{\left.\frac{dy}{d\theta}\right|_{y^\circ}} = V(\theta)$ = How parameter affects dist'n at y°

Prior: $\pi(\theta)d\theta = |V(\theta)|\,d\theta$

FRMY 2010

Ex: $y = X\beta + \sigma z$ $\qquad V(\theta) = (X, z^\circ(\theta))$ $\qquad z^\circ(\theta) = (y^\circ - X\beta)/\sigma$

Prior: $\boxed{\pi(\theta)d\theta = d\beta\, d\sigma/\sigma}$ $\qquad$ (Jeffreys original $d\beta\, d\sigma/\sigma^p$)

(ii) Often can write $\hat\theta = \hat\theta(\theta, \hat z)$

Continuity $\Rightarrow$ $\underset{p\times p}{\left.\frac{d\hat\theta}{d\theta}\right|_{\hat\theta^\circ}} = \bar V(\theta)$ $\qquad$ Prior: $\pi(\theta)d\theta = |\bar V(\theta)|\,d\theta$

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\}\, h(s)$$

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\} h(s)$$

Curious properties: 1) Expd Info = Obs Info. 2) $J_{\varphi\varphi}(\varphi)$ free of s.

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\}\, h(s)$$

Curious properties: 1) Expd Info = Obs Info. 2) $j_{\varphi\varphi}(\varphi)$ free of s.

3) Jeffreys prior $= |i_{\varphi\varphi}(\varphi)|^{1/2} = |j_{\varphi\varphi}(\varphi)|^{1/2}$ can be "Badly biased" if $p > 1$

Jeffreys 1961

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\}\, h(s)$$

Curious properties: 1) Expd Info = Obs Info. 2) $j_{\varphi\varphi}(\varphi)$ free of s.

3) Jeffreys prior $= |i_{\varphi\varphi}(\varphi)|^{1/2} = |j_{\varphi\varphi}(\varphi)|^{1/2}$ can be "Badly biased" if $p > 1$

Jeffreys 1961

Another property (recent)

4) Prior with repetition validity exists (2nd order, moderate deviations)

Write: $\varphi = \hat{\varphi}^0 + \rho\, \underset{\sim}{u}$ $\underset{\sim}{u}$ is unit vector

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\}\, h(s)$$

Curious properties: 1) Expd Info = Obs Info. 2) $j_{\varphi\varphi}(\varphi)$ free of s.

3) Jeffreys prior $= |i_{\varphi\varphi}(\varphi)|^{1/2} = |j_{\varphi\varphi}(\varphi)|^{1/2}$ can be "Badly biased" if $p > 1$

Jeffreys 1961

Another property (recent)

4) Prior with "repetition validity" exists (2nd order, moderate deviations)

Write: $\varphi = \hat{\varphi}^0 + \rho\, \underset{\sim}{u}$ $\underset{\sim}{u}$ is unit vector

2nd order prior $= |J_{\rho\rho}(\varphi)|^{1/2}$ one dimensional ... radially

(g) Priors from information

Case: Exponential models; e.g. GLM canonical link

$$f(s;\varphi) = \exp\{\varphi' s - K(\varphi)\}\, h(s)$$

Curious properties: 1) Expd Info = Obs Info. 2) $j_{\varphi\varphi}(\varphi)$ free of s.

3) Jeffreys prior $= |i_{\varphi\varphi}(\varphi)|^{1/2} = |j_{\varphi\varphi}(\varphi)|^{1/2}$ can be "Badly biased" if $p > 1$

Jeffreys 1961

Another property (recent)

4) Prior with "repetition validity" exists (2nd order, moderate deviations)

Write: $\varphi = \hat{\varphi}^0 + \rho\, \underset{\sim}{u}$ $\underset{\sim}{u}$ is unit vector

2nd order prior $= |j_{ee}(\varphi)|^{1/2}$ one dimensional ... radially

Uniqueness ... 2nd order

Scalar φ: Welch Peers 1963. Vector φ: F 2012 JRSSB in review

(h) Measure the "bias" from a prior

1) The simple Normal $y \sim N(\theta; 1)$

Try prior $\pi(\theta) = \exp\{a\theta\}$

(h) Measure the "bias" from a prior

1) The simple Normal $y \sim N(\theta; 1)$

Try prior $\pi(\theta) = \exp\{a\theta\}$

Get: $\pi(\theta; y)$ $\theta \sim N(y^\circ + a; 1)$

Bias from prior/Shift of Stdzed Lik. is $B = a$

(h) Measure the "bias" from a prior

1) The simple Normal $y \sim N(\theta; 1)$

Try prior $\pi(\theta) = \exp\{a\theta\}$

Get: $\pi(\theta; y)$ $\theta \sim N(y^\circ + a; 1)$

Bias from prior/Shift of Stdzed Lik. is $B = a$

2) First order analysis

Stdzed $\log L(\theta) = -\frac{1}{2}(\theta - \hat{\theta})' j_{\theta\theta}(\hat{\theta})(\theta - \hat{\theta})$ --- $O(n^{-1/2})$

Stdzed prior $\pi(\theta) = \exp\{a'\theta/n^{1/2}\}$

(h) Measure the "bias" from a prior

1) The simple Normal $y \sim N(\theta; 1)$

Try prior $\pi(\theta) = \exp\{a\theta\}$

Get: $\pi(\theta; y)$ $\quad \theta \sim N(y^\circ + a; 1)$

Bias from prior/Shift of Stdzed Lik. is $B = a$

2) First order analysis

Stdzed $\log L(\theta) = -\frac{1}{2}(\theta - \hat\theta)' j_{\theta\theta}(\hat\theta)(\theta - \hat\theta)$ $\quad \ldots \quad O(n^{-1/2})$

Stdzed prior $\pi(\theta) = \exp\{a'\theta/n^{1/2}\}$

Get: $\pi(\theta; y)$ ... just "log Likelihood"

... no effect from prior

3) Measure the bias from a prior: 2nd order analysis

Case: Scalar parameter (vector analogous)

3) Measure the bias from a prior: 2nd order analysis

Case: Scalar parameter (vector analogous)

a) Welch-Peers location parameter 1963

$$\beta = \int_{\hat{\varphi}^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi) \cdot d\varphi \qquad "d\beta = j^{1/2} d\varphi"$$

3) Measure the bias from a prior: 2nd order analysis

Case: Scalar parameter (vector analogous)

a) Welch-Peers location parameter

$$\beta = \int_{\hat{\varphi}^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi) \cdot d\varphi \qquad "d\beta = j^{1/2} d\varphi"$$

β has constant info = 1

3) Measure the bias from a prior: 2nd order analysis

Case: Scalar parameter (vector analogous)

a) Welch-Peers location parameter

$$\beta = \int_{\hat{\varphi}^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi) \cdot d\varphi \qquad "d\beta = j^{1/2} d\varphi"$$

β has constant info = 1

"flat" re β $\Rightarrow$ "Bayes = frequency"

3) Measure the bias from a prior : 2nd order analysis

Case : Scalar parameter (vector analogous)

a) Welch-Peers location parameter

$$\beta = \int_{\hat{\varphi}^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi) \cdot d\varphi \qquad \text{"} d\beta = j^{1/2} d\varphi \text{"}$$

β has constant info = 1

"flat" re β $\Rightarrow$ "Bayes = frequency"

b) Proposed prior $\pi(\varphi)$ re parameter φ

equiv to prior $\pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ re parameter β

3) Measure the bias from a prior : 2nd order analysis

Case : Scalar parameter (vector analogous)

a) Welch-Peers location parameter

$$\beta = \int_{\hat{\varphi}^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi) \cdot d\varphi \qquad \text{"} d\beta = j^{1/2} d\varphi \text{"}$$

| β has constant info = 1

"flat" re β $\Rightarrow$ "Bayes = frequency"

b) Proposed prior $\pi(\varphi)$ re parameter φ

equiv to prior $\pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ re parameter β

c) Bayes = freq'cy requires $\pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ be flat re β

3) Measure the bias from a prior: 2nd order analysis

Case: Scalar parameter (vector analogous)

a) Welch-Peers location parameter

$$\beta = \int_{\hat\varphi^0}^{\varphi} J_{\varphi\varphi}^{1/2}(\varphi)\cdot d\varphi \qquad "d\beta = j^{1/2} d\varphi"$$

| β has constant info = 1

"flat" re β $\Rightarrow$ "Bayes = frequency"

b) Proposed prior $\pi(\varphi)$ re parameter φ

equiv to prior $\pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ re parameter β

c) Bayes = freq'cy requires $\pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ be flat re β

d) % Bias $B = \frac{\partial}{\partial\beta} \log \pi(\varphi)\, j_{\varphi\varphi}^{-1/2}(\varphi)$ $\qquad r \longrightarrow r + B$

$$= \left\{ \frac{\partial}{\partial\varphi} \log \pi(\theta) - \frac{\partial}{\partial\varphi} \log J_{\varphi\varphi}^{1/2}(\varphi) \right\}\Big|_{\hat\varphi^0} \cdot J_{\varphi\varphi}^{-1/2}(\hat\varphi^0)$$

(given) − (Jeffreys)

= Amount by which stdgd likelihood is displaced

F 2012 PJSOR To appear

References:

Jeffreys 1946 Proc. R.S. A An invariant form for the prior probability ...

Welch Peers 1963 JRSS B On formula for confidence points - -

David Stone Zidek 1973 JRSSB Marginalization Paradoxes ...

Bernardo 1979 JRSSB Reference posterior distributions xxx

F Reid 2002 JSPI Strong matching for frequentist and Bayesian ... 211

F Reid Marras Yi 2010 JRSSB Default priors for Bayesian and freq 239

F F F 2010 Stat Res (Efron Vol) Parameter curvature revisited and the B-f 249

F 2011 Stat. S. Res Is Bayes posterior just quick and dirty confidence 247

F 2012 PJSOR 2012 Bias in Bayes and how to measure it 261

F 2012 JRSSB (in review) An adjusted Jeffreys prior 265

Can find the good prior - via continuity: $|V(\theta)|$ 239

- via info: $J_{ec}^{1/2}(\theta)$ 265

If not good, can calculate bias: B 261