

Chapter 10. Probabilistic Reasoning

Likan Zhan

Beijing Language and Culture University

2020-02-14

<https://likan.info>

zhanlikan@blcu.edu.cn

Contents

- “Person-Who” Statistics
- Probabilistic Reasoning and the Misunderstanding of Psychology
- Psychological Research on Probabilistic Reasoning
 - Insufficient Use of Probabilistic Information
 - Failure to Use Sample-Size Information
 - The Gambler’s Fallacy
 - A Further Word About Statistics and Probability

Probabilistic Reasoning

Probabilistic Reasoning

- Men are taller than women, right?

Probabilistic Reasoning

- Men are taller than women, right?
- All men are taller than all women, right?

Brad Pitt and Angelina Jolie



People seem to be in a land of sometimes and perhaps, and they had hoped to go on living with always and with certainty.

– Bronowski, 1978

Person-Who Statistics



Probabilistic Reasoning

Probabilistic Reasoning

- At a particular dosage level, 57 percent of patients taking this medication receive relief in two hours.

Probabilistic Reasoning

- At a particular dosage level, 57 percent of patients taking this medication receive relief in two hours.
- Unless he or she exercises and changes diet, he or she has a high risk of heart attack.

Probabilistic Reasoning

- At a particular dosage level, 57 percent of patients taking this medication receive relief in two hours.
- Unless he or she exercises and changes diet, he or she has a high risk of heart attack.
- There is an 80 percent probability of a magnitude 8.0 or greater earthquake in a certain area in the next 30 years.

Daniel Kahneman



Daniel Kahneman

Prize share: 1/2

Half of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2002 was awarded to Daniel Kahneman *for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty.*

Insufficient Use of Probabilistic Information

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);
- 再假设有一种检查可以百分百地诊断出真的携带该病毒的人;

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);
- 再假设有一种检查可以百分百地诊断出真的携带该病毒的人;
- 最后, 假设这个检查有 5% 的可能性把没有携带者说成有。也就是说这项检查再没有携带 HIV 的人中, 也会错误的检测出 5% 的人是病毒携带者.

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);
- 再假设有一种检查可以百分百地诊断出真的携带该病毒的人;
- 最后, 假设这个检查有 5% 的可能性把没有携带者说成有。也就是说这项检查再没有携带 HIV 的人中, 也会错误的检测出 5% 的人是病毒携带者.
- 假设我们随便找一个人来进行这项检查, 得到了阳性反应, 亦即此认为 HIV 携带者.

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);
- 再假设有一种检查可以百分百地诊断出真的携带该病毒的人;
- 最后, 假设这个检查有 5% 的可能性把没有携带者说成有。也就是说这项检查再没有携带 HIV 的人中, 也会错误的检测出 5% 的人是病毒携带者.
- 假设我们随便找一个人来进行这项检查, 得到了阳性反应, 亦即此认为 HIV 携带者.
- 那么他真的是 HIV 携带者的概率是多少呢?
A.95%; B. 接近 2%!

Insufficient Use of Probabilistic Information

- 如果每 1000 个人中有一个人携带艾滋病病毒 (HIV);
- 再假设有一种检查可以百分百地诊断出真的携带该病毒的人;
- 最后, 假设这个检查有 5% 的可能性把没有携带者说成有。也就是说这项检查再没有携带 HIV 的人中, 也会错误的检测出 5% 的人是病毒携带者.
- 假设我们随便找一个人来进行这项检查, 得到了阳性反应, 亦即此认为 HIV 携带者.
- 那么他真的是 HIV 携带者的概率是多少呢?
A. 95%; B. 接近 2%!

Failure to Use Sample-Size Information

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。
- 正如大家所知道的, 大约有 50% 的婴儿是男孩。

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。
- 正如大家所知道的, 大约有 50% 的婴儿是男孩。
- 当然, 真正的百分比每天都不一样, 有时候高于 50%, 有时候低于 50%。

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。
- 正如大家所知道的, 大约有 50% 的婴儿是男孩。
- 当然, 真正的百分比每天都不一样, 有时候高于 50%, 有时候低于 50%。
- 在一年的时期内, 每一所医院都记录了出生的男婴比例高于 60% 的天数。

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。
- 正如大家所知道的, 大约有 50% 的婴儿是男孩。
- 当然, 真正的百分比每天都不一样, 有时候高于 50%, 有时候低于 50%。
- 在一年的时期内, 每一所医院都记录了出生的男婴比例高于 60% 的天数。
- 你认为哪一个医院纪录的天数多?
A. 大医院; B. 小医院; C. 基本一样

Failure to Use Sample-Size Information

- 一个小镇有大小两所医院。
- 在大医院里每天大约有 45 个婴儿出生, 在小医院里每天大约有 15 个婴儿出生。
- 正如大家所知道的, 大约有 50% 的婴儿是男孩。
- 当然, 真正的百分比每天都不一样, 有时候高于 50%, 有时候低于 50%。
- 在一年的时期内, 每一所医院都记录了出生的男婴比例高于 60% 的天数。
- 你认为哪一个医院纪录的天数多?
A. 大医院; B. 小医院; C. 基本一样

Failure to Use Sample-Size Information

Failure to Use Sample-Size Information

- 假设一个缸里装满了球, 其中一个有 $2/3$ 是一种颜色, $1/3$ 是另一种颜色.

Failure to Use Sample-Size Information

- 假设一个缸里装满了球, 其中一个有 $2/3$ 是一种颜色, $1/3$ 是另一种颜色.
- 一个人从缸里拿出 5 个球发现 4 个是红色的, 1 个是白色的.

Failure to Use Sample-Size Information

- 假设一个缸里装满了球, 其中一个有 $2/3$ 是一种颜色, $1/3$ 是另一种颜色.
- 一个人从缸里拿出 5 个球发现 4 个是红色的, 1 个是白色的.
- 另一个人从里面拿出 20 个球, 发现有 12 个是红色的, 8 个是白色的.

Failure to Use Sample-Size Information

- 假设一个缸里装满了球, 其中一个有 $2/3$ 是一种颜色, $1/3$ 是另一种颜色.
- 一个人从缸里拿出 5 个球发现 4 个是红色的, 1 个是白色的.
- 另一个人从里面拿出 20 个球, 发现有 12 个是红色的, 8 个是白色的.
- 哪一个人更有自信的认为这个缸里有 $2/3$ 的球是红色的, 有 $1/3$ 的球是白色的, 而不是有 $1/3$ 的球是红色的, 有 $2/3$ 的球是白色的? 这两个人会给出什么样的概率来说明这两种说法的正确程度呢?

Failure to Use Sample-Size Information

Failure to Use Sample-Size Information

- A study of 3,141 counties in the United States found that

Failure to Use Sample-Size Information

- A study of 3,141 counties in the United States found that
- The counties in which the incidence of kidney cancer was lowest tended to be rural counties that were sparsely populated.

Failure to Use Sample-Size Information

- A study of 3,141 counties in the United States found that
- The counties in which the incidence of kidney cancer was lowest tended to be rural counties that were sparsely populated.
- The counties in which the incidence of kidney cancer was highest tended to be rural counties that were sparsely populated!

Failure to Use Sample-Size Information

Failure to Use Sample-Size Information

- Your physician is making assumptions about your entire composition from a tiny sample.

Failure to Use Sample-Size Information

- Your physician is making assumptions about your entire composition from a tiny sample.
- We often are taking a small sample to represent a much larger population of behavior.

The Gambler's Fallacy

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?
 - b. It is more likely that heads will come up than tails?

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?
 - b. It is more likely that heads will come up than tails?
 - c. Heads and tails are equally probable on the sixth toss?

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?
 - b. It is more likely that heads will come up than tails?
 - c. Heads and tails are equally probable on the sixth toss?

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?
 - b. It is more likely that heads will come up than tails?
 - c. Heads and tails are equally probable on the sixth toss?
- Problem B: When playing slot machines, people win something one out of every 10 times. Julie, however, has just won on her first four plays. What are her chances of winning the next time she plays?

The Gambler's Fallacy

- Problem A: Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads five times in a row. For the sixth toss, do you think that
 - a. It is more likely that tails will come up than heads?
 - b. It is more likely that heads will come up than tails?
 - c. Heads and tails are equally probable on the sixth toss?
- Problem B: When playing slot machines, people win something one out of every 10 times. Julie, however, has just won on her first four plays. What are her chances of winning the next time she plays? One out of 10

The Gambler's Fallacy

- **The gambler's fallacy**: the tendency for people to see links between events in the past and events in the future when the two are really independent.

Randomrization



Questions?