The Association of Mathematics Educators
The Singapore Mathematical Society

Conference for Mathematics Teachers

Workshop: Probability Through Games

6 June 2013

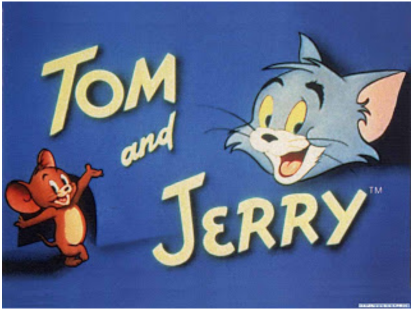
David Chew

Department of Statistics & Applied Probability
National University of Singapore

Quiz: how good are you at probability?

Coin Toss: Tom and Jerry

Tom and Jerry were instructed to toss a fair coin 200 times and record the outcomes of the tosses. One faithfully tossed the coin and recorded the outcomes. One cheated and did not toss at all, he simply made up the outcomes.



Quiz: how good are you at probability?

Coin Toss: Tom and Jerry

Tom's sequence:

```
THHTT TTHTH HHTTT HTTHT HHTTH TTTTH HHHTH HHHTH THHTT HTTTT
TTHTT HTHTT HTTHT HTHHT THTTH HHTTT THTTH HHHTH THTHT HHTHT
HHHHH HTTHT HHHHT TTHTH HTTTH TTHTH TTTHH THTHT TTHHH HHTTT
HHHTH HTHHT HTTHT HTTTH HHHHT THHTT HHHTH THTTT HTHHT TTHHH
```

Jerry's sequence:

```
TTTHT HHHTH THTHT TTTHT TTHTH HTTHH THTHH TTHTT HHHTH HTHHT
THHHH THTTT HTTTT THHTH HHTTH THTTH THTTH HHTTT HHTTH TTTTT
THTHT HHHHT HHHTH THTHT HHHHH THTHT HHTHH HHHTH HHHTH HTHTT
THTTT HTHHT TTHHH HTHHH HTTHH THTTH THHTT HHHTH HTHHH HHHTH
```

Who is likely to cheat? How can you find out?

Quiz: how good are you at probability?

Lie Detector tests

A recent report of the US National Academy of Sciences examines the evidence for the effectiveness of lie detector tests in large scale screening of employees and as a part of terrorist counter measures.

`http://www.nap.edu/books/0309084369/html/`

Quiz: how good are you at probability?

Lie Detector tests

- Earlier data provided by noted legal statistician Joe Gastwirth is as follows. A positive reading on the test indicates that the subject is lying.
- Probability of positive reading in subjects who are lying is 88%.
- Probability of negative reading in subjects who are telling the truth is 86%.
- In security screening applications it is assumed that 99% of individuals being screened are telling the truth.

Quiz: how good are you at probability?

Lie Detector tests

A subject produces a positive reading on the test. What is the probability that the test is incorrect and the subject is telling the truth?

(a) Less than 10%.
(b) At least 10% and less than 20%.
(c) At least 20% and less than 50%.
(d) At least 50% and less than 90%.
(e) At least 90%.

Quiz: how good are you at probability?

The Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats.

You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?"

Is it to your advantage to switch your choice?

Outline

1. Introduction: What is probability?
2. Different approaches in assigning probability
    2.1 Classical approach:
    Probability based on a *known* system of generating the outcomes
    2.2 Frequentist approach:
    Probability based on *long-run frequency*
    2.3 Subjective approach:
    Probability based on *subjective belief*
3. Consistency: Rules of Probability
4. Independence
5. Law of Total Probability
6. Baye's Rule
7. Simple Random Sample and Surveys

1. Introduction: What is probability?

What does

“the probability of getting a six in a roll of die is 1/6”

mean?

2. Different approaches in assigning probability

2.1 Classical approach

- Probability based on a known system of generating the outcomes (though random), such as tossing a coin, rolling a die etc...

- Toss a fair coin once. What is the probability of getting a head?
  - We have no reason to believe $P(\text{Head}) \neq P(\text{Tail})$.
  - That is, $P(\text{Head}) = P(\text{Tail}) = p$.
  - Since $P(\text{Head}) + P(\text{Tail}) = 1$, then $2p = 1$. Hence

$$P(\text{Head}) = P(\text{Tail}) = \frac{1}{2}.$$

2. Different approaches in assigning probability

2.1 Classical approach

- Toss this coin independently three times. What is the probability of getting 2 heads and 1 tail?
  - We have no reason for one of the outcomes listed below to have a probability different from the rest.

HHH, HHT, HTH, HTT, THH, THT, TTH,

  - 3 out of 8 possible outcomes gives "2 heads and 1 tail", so $P(\text{2 heads and 1 tail}) =$.
- **General Result:** For an experiment with many outcomes, all outcomes are assumed to be equally likely, then

$$\boxed{\text{Prob (A)} = \frac{\text{No. of ways outcome } A \text{ can occur}}{\text{Total no. of possible outcomes}}}$$

Activity I: Flipping Coins

- Each one of you pick up a coin.
- Flip it 200 times, recording the outcome each time you flip the coin.
- Count the number of heads in the first

$$10, 20, 30, 40, 50, \ldots, 190, 200 \text{ flips.}$$

- Divide the first number by 10, the second by 20, . . . , the last by 200.
- Plot those numbers vs $10, 20, 30, 40, 50, \ldots, 190, 200$.
- What do you observe?

2. Different approaches in assigning probability

2.2 Frequentist approach

- How do you know if a coin, used in a gambling den, is fair?
- We can find out by experiment: many repeated flips.
- Flip this coin 10 times and record the frequency of heads:

10 flips	T	H	T	T	H	H	T	T	T	T
frequency	0	0.5	0.333	0.25	0.4	0.5	0.429	0.375	0.333	0.3

- Flip this coin

2. Different approaches in assigning probability

2.2 Frequentist approach: A Real Story . . .

- English mathematician John Kerrich was in Copenhagen when World War II broke out and he was imprisoned by the Nazis.

- To help pass the time he tossed a coin 10,000 times and recorded the results in order to look for any evidence of bias in his coin (there wasn't any), a good illustration of the frequency interpretation of probability.

2. Different approaches in assigning probability

2.2 Frequentist approach: Tom and Jerry

What about Tom and Jerry's plots?

What is your view now?

2. Different approaches in assigning probability

2.3 Subjective approach

- **Question:**
  What is the probability of you getting an A in your mathematics examination?
- This is a unique situation, unlikely to be repeated many, many times.
- Need to assign a subjective (or personal) probability.

2. Different approaches in assigning probability

2.3 Subjective approach: Airplane Collisions

- The following report was given in *The West Australian* newspaper following an accident where two jumbo jets collided on a runway:
  - "Mr Webster Todd, Chairman of the American National Transportation Safety Board, said today that statistics showed that the chances of two jumbo jets colliding on the ground were about 6 million to one."
- Terry Speed, a statistician at University of Western Australia, wrote to the chairman to ask how this value was calculated. He received the following reply.

2. Different approaches in assigning probability

2.3 Subjective approach: Airplane Collisions

Dear Professor Speed,

In response to your aerogram of April 5, 1977, the chairman's statement concerning the chances of two jumbo jets colliding (6 million to one) has no statistical validity, nor was it intended to be a rigorous or precise probability statement. The statement was made to emphasize the intuitive feeling that such an occurrence indeed has a very remote but not impossible chance of happening.

Thank you for your interest in this regard.

Whatever approach you use, consistency is required.

Rule 1 (Convexity):

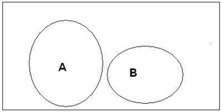
$$\boxed{0 \leq P(A) \leq 1}.$$

- If we know A is impossible, then $P(A) = 0$;
- If we know A is always true, then

3. Rules of Probability

Rule 2: Addition

- Two events A and B are said to be mutually-exclusive if A and B have no common outcomes.



- Rule 2 (Addition):
  When A and B are two mutually-exclusive events, then

$$\boxed{P(A \text{ or } B) = P(A) + P(B)}.$$

- Derived Rule 1: $P(A^c) = 1 - P(A)$
- Derived Rule 2: Bigger event ($A \subseteq B$), bigger probability

$$P(A) \leq P(B)$$

3. Rules of Probability

Rule 2: Addition: About Linda

Consider the following description of Linda:

> *Linda is 31 years old, single, outspoken, and very bright. She majored in Philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.*

Which of the following alternatives is more probable? And why?

(a) Linda is a bank teller.

(b) Linda is a bank teller and active in feminist movement.

Conditional Probability

Roll two fair dice.

- What is the probability that the sum of the 2 dice is an even number?
- Given that the first die shows a 5, what is the probability that the sum of the 2 dice is an even number?

Conditional Probability

The possible outcomes on two dice:

		2nd die					
		1	2	3	4	5	6
	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
1st die	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditional Probability

Let B denote the event that the sum of the 2 dice is an even number, and A the event that the first die is a 5.

- $P(B) =$

- $P(B|A) =$

Conditional Probability

Definition

Let A and B be two events.

If $P(A) > 0$, the **conditional probability of B given A** is defined as

$$P(B|A) = \frac{P(AB)}{P(A)}.$$

3. Rules of Probability

Rule 3: Multiplication

- Rule 3 (Multiplication):
  To find the probability that A and B occur, we have

$$P(A \text{ and } B) = P(A)P(B|A).$$

- Interpretation of the right hand side. Calculate the probability
  - Calculate $P(A)$, i.e, A happens first;
  - Calculate the probability of B occurring given the occurrence of A, i.e., $P(B|A)$.

- Similarly,

$$P(A \text{ and } B) = P(B)P(A|B)$$

4. Independence

Independence of 2 events

- A special, but very important, case in which information that B is true does nothing to change our uncertainty about A (and vice versa).
- That is, $P(A|B) = P(A)$.
- In this case, we say that A and B are independent.
- Mathematically, we have

$$\boxed{P(A \text{ and } B) = P(A) \times P(B)}.$$

- Reasoning:

$$\begin{aligned} P(A \text{ and } B) &= P(A|B) \times P(B) \qquad \text{(Multiplication rule)} \\ &= P(A) \times P(B). \end{aligned}$$

Events A, B and C are independent if

1. Any two of them are

2. $P(A \text{ and } B \text{ and } C) = P(A)P(B)P(C)$.

4. Independence

Misunderstanding Independence: People v. Collins

- In 1964 Mrs Juanita Brooks was knocked over while walking home with her shopping basket. When she got up, she saw a young woman running away and found that her purse was missing.
- The young woman was described as having blond hair in a pony tail and wearing something dark.
- Another witness saw such a woman get into a yellow car driven by a black male with a beard and a moustache.
- Collins and his wife Janet fitted the description.

The following probabilities were presented by the prosecution concerning characteristics of the suspects (no evidence was presented to justify the numerical values given):

Yellow Car	1/10	Girl with blond hair	1/3
Man with moustache	1/4	Black man with beard	1/10
Girl with ponytail	1/10	Interracial couple in car	1/1000

4. Independence

Misunderstanding Independence: People v. Collins

- To help with a weak identification of the suspect in a lineup, the prosecution called on a college mathematics instructor.
- The instructor explained the rule for multiplying probabilities for independent events, and multiplying the probabilities in the table concluded that the chance of a random couple meeting all the characteristics described was about
- According to the prosecutor, this is the chance "that any other couple possessed the distinctive characteristics of the defendants."
- The jury convicted.
- Anything wrong with this reasoning?

4. Independence

Misunderstanding Independence: People v. Collins

-

- For example, “a man with a mustache” and “black man with a beard”. Most men with beards also have mustaches.
- Furthermore, if you have a “girl with blond hair” and “negro man with beard”, the chance of having an “inter-racial couple” is close to 1 not 1/1000, so that from this alone the answer is too small by a factor of about 1000.
- Also, the prosecution had presented no evidence to support the values chosen for their probabilities.
- HAPPY ENDING:
  The verdict was overturned by the California Supreme Court.

4. Independence

Misunderstanding Independence: Sally Clark & Lucy De Berk

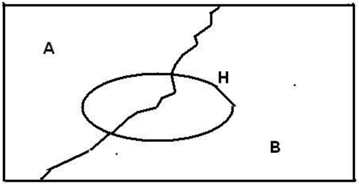
- Another legal case where an inappropriate assumption of independence played a role in statistical evidence presented was the Sally Clark case in the UK.
- Sally Clark was convicted of murdering her two infant children, partly based on the reasoning that it was very unlikely that two children in the same family would both die from Sudden Infant Death Syndrome. (In calculations presented by an expert witness for the prosecution the deaths were assumed independent in calculating the probability of this happening by chance).
- See `http://en.wikipedia.org/wiki/Sally_Clark`
- A more recent case is that of Dutch nurse Lucy De Berk, who was jailed for life in 2003 for the murders of seven patients and the attempted murders of another three.
- See `http://news.bbc.co.uk/2/hi/europe/8620997.stm`

5. Law of Total Probability

- If A and B are mutually-exclusive and exhaustive events, then for any event H,

$$\boxed{P(H) = P(H|A)P(A) + P(H|B)P(B)}.$$

Note: A and B are mutually exhaustive if $A = B^C$.



- Reasoning:

$$\begin{aligned} P(H) &= P(H \text{ and } A) + P(H \text{ and } B) && \text{(Rule 2)} \\ &= P(H|A)P(A) + P(H|B)P(B) && \text{(Rule 3)} \end{aligned}$$

5. Law of Total Probability

Example

Fred is playing in a chess tournament. He has just won his match in the first round. In the second round he is due to play either Bernard or Tessa depending on which of those two wins their first match. Knowing what he does about their respective abilities, he assigns $\frac{3}{4}$ to the probability that he will have to play Bernard, $P(B)$ and $\frac{1}{4}$ to the probability that he will have to play Tessa, $P(T)$. He further assigns $\frac{1}{4}$ to probability that he will beat Bernard if he has to play him, $P(W|B)$, and $\frac{1}{2}$ to probability that he will beat Tessa if he has to play her, $P(W|T)$. What is the probability that Fred will win his second match?

Solution:
Note that the events B and T are mutually-exclusive and exhaustive. Then the probability that Fred wins his next match is

Activity 2: The Monty Hall Problem

With acknowledgement: *The MythBusters*

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats.

You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?"

Is it to your advantage to switch your choice? Let's play the game to find out.

or

?

6. Bayes' Rule

If A and B are mutually-exclusive and exhaustive events, then for any event H,

$$
\begin{aligned}
P(A|H) &= \frac{P(AH)}{P(H)} \\
&= \frac{P(H|A)P(A)}{P(H|A)P(A) + P(H|B)P(B)}.
\end{aligned}
$$

Here we used the Multiplication rule for the numerator and the Law of Total Probability for the denominator.

6. Bayes' Rule

Lie Detector Tests: Solution

Let T denote the event that the subject is telling the truth and R the event that the subject produces a positive reading on the test. Then what we are seeking is $P(T|R)$.

The information given means that

$$P(T) = 0.99, \quad P(R|T) = 0.14, \quad P(R|T^c) = 0.88.$$

It then follows that

$$P(T|R) =$$

Activity (III): Chocolate sampling

With acknowledgement: Dr Alex Cook

- There are 2 bags of chocolates (with the same contents). Divide yourselves into 2 groups. Each group takes a bag.
- Each of you take turn to draw out 7 pieces of chocolates.
  - Record the total weight of those 7 pieces of chocolates you have drawn.
  - The average of those 7 pieces will be used to estimate the average weight of chocolates in that bag.
  - Return the chocolates to the bag. The next person draws.
- The one with the closest estimate (within 2 grams of the correct answer) wins the bag.
- Did you manage to get a good estimate? If not, why?

7. Simple Random Sample and Surveys

The 1936 US Presidential Election

- USA is struggling to recover from Great Depression.
- Candidates: Roosevelt vs Landon
- Most thought Roosevelt would be an easy winner.
- The Literary Digest, which had correctly predicted the winner of the last 5 elections, announced that Landon would be the winner, with 57% of the popular vote.
- The prediction was based on a survey to which 2.4 million people replied.

7. Simple Random Sample and Surveys

The Results

In the end, Roosevelt won with 62% of the votes.

What went wrong?

7. Simple Random Sample and Surveys

The Predictions

	Sample size	Roosevelt	Landon
Literary Digest	2,400,000	43%	57%
George Gallup	50,000	56%	44%
Actual	–	62%	38%

7. Simple Random Sample and Surveys

What Went Wrong

The Digest mailed questionnaires to 10 million people.

- Names are taken from telephone books and club membership lists. This tend to screen out the poor.

  When a selection procedure is biased, taking a large sample does not help. This just repeats the basic mistake on a larger scale.

- The response rate was 24%.

  Non-respondents can be very different from respondents. When there is a high non-response rate, look out for non-response bias.

7. Simple Random Sample and Surveys

Probability Methods for Sampling

- The interviewers have no discretion as to who to interview.
- There is a definite procedure for selecting the sample, and it involves the use of chance.

7. Simple Random Sample and Surveys

Using Chance in Surveys

- Even in 1948, a few survey organizations using probability methods to draw their samples.
- Now, many organizations do.
- When you come across results from a survey the next time, ask yourself: how was the sample taken?
  (Stand at Orchard road and conduct interview?)

Summary

- Probability can sometimes be counter-intuitive.
- It has useful applications in a variety of (real life) problems.
- It has to be handled with care.