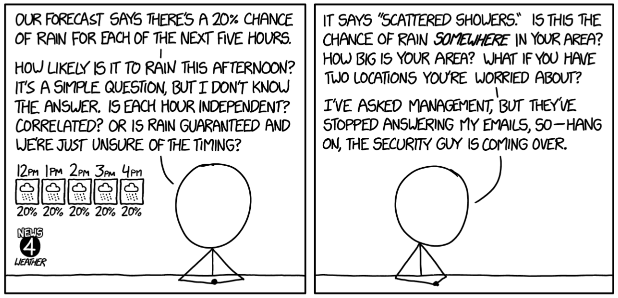
Introduction to Probability and Statistics Stat 20 Fall 2018



https://imgs.xkcd.com/comics/meteorologist.png

Lecture 4: 9/4/2018

Chapters 16.1,16.2, 13,14

Shobhana M. Stoyanov

Warm-up: 16.R.10(a)

- 200 draws are made *at random with replacement* from the box :

-3	-2	-1	0	1	2	3

- If the sum of the numbers drawn is 30, what is their average?

A. 0.3 B. 0 C. 0.15

Warm-up: 16.R.10(b)

- 200 draws are made at random with replacement from the box :

-3	-2	-1	0	1	2	3

There are 2 alternatives:

(i)winning $1 if the **sum** of the 200 numbers drawn is between -5 and 5; or

(ii) winning $1 if the **average** of the 200 numbers drawn is between -0.025 and +0.025. Which is better?

A. (i) B. (ii) C. Both same

Chapter 16 : The Law of Averages

- John Kerrich, a South African mathematician was interned in Denmark during WWII.
- To pass the time, he experimented with coin tossing, recording his findings.
- Question: What does the *law of averages* say?
- Quote in text: *The roulette wheel has neither conscience nor memory. (Joseph Bertrand)*
- The number of heads will be around *half* the number of tosses, give or take some error.

Simulating Kerrich's experiments

- heads=sample(x=c(0,1),size=n,replace=TRUE)
- count_heads=cumsum(heads)
- ## cumsum: **cum**ulative **sum**
- exp_heads=(1:n)/2
- diff_heads=count_heads-exp_heads
- percent_heads=(count_heads/(1:n))*100

9/3/18 5

Law of Averages

- Notice that as the number of tosses of a fair coin increases, the *observed error* (number of heads – half the number of tosses) increases.
- The *percentage* of heads observed comes very close to 50%
- *Law of averages*: The long run proportion of heads is very close to 50%.

9/3/18 6

Chance processes

- When we perform some action whose outcome is random, each time we perform the action, the outcome will be different.
- Example: Take random sample of voters, count immigration ban supporters. Do it again. And again.
- To what extent does chance influence our results? This is our subject of investigation of the next few chapters.

9/3/18 7

Chance processes

- We will:
  1. Find an analogy between the process we are studying and the process of drawing numbered tickets at random from a box.
  2. Connect the variability that we are interested in to the chance variability in the sum of the numbers drawn from the box.
- We use a *BOX MODEL* to analyze the chance processes.

9/3/18 8

Example

- A gambler loses 10 times running at roulette. He decides to continue playing because he is due for a win, by the law of averages. A bystander advises him to quit, on the grounds that his luck is cold. Who is right?

A. Gambler B. Bystander C. Neither

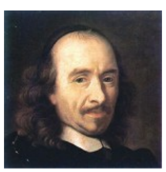
Chapter 13: Chance

- How would you define *chance*?
  - Toss a coin: what is the chance that you see heads?
- Frequency theory or *frequentist* approach to chance
  - This works for processes that can be repeated as often as desired, independently, and under the same conditions.

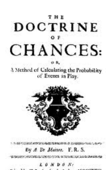
The *chance* or the *probability* of something happening is the percentage of time that it is expected to happen when the basic process is repeated over and over again, independently, and under the same conditions (long run relative frequency).

Origins of probability…murky?

Questions that arose from gambling with dice.







Rules governing chance

- Chances (probabilities) are between 0 (0%) and 1 (100%). Remember, they are *proportions*.
- The process is called a *random experiment*. We know all the possible outcomes, but are uncertain about which will occur on a particular *trial*.
- A collection of all possible outcomes of an action is called a *sample space*. Usually denoted by S or Ω.
  - Example: Toss a coin twice. Sample space = {HH, HT, TH, TT}
- An *event* is a collection of outcomes. Usually denoted by A, B etc.
  - Example: S as defined above. Let A be the event of no heads. A={TT}

Rules governing chance

- If an event A is *certain* to occur, we say the chance of A (P(A)) is 100%.
- If an event A is *impossible*, we say P(A) =0%.
- The *opposite event* of A is called the *complement* of A and is denoted by A^C or "not A".
- $P(A^C)$ = P(not A) = 100% - P(A) = 1 – P(A)
- P(A) + $P(A^C)$ = 1 = 100%

9/3/18 13

Rules governing chance

- If all the possible outcomes are *equally likely*, then each outcome has probability 1/n, where *n* = number of possible outcomes.
- If an event A contains k possible outcomes,

  then P(A) = *k/n*.
- Outcomes don't have to be equally likely. Consider the lottery: P(win) is not the same as P(don't win)

9/3/18 14

Examples: Die rolls

1. Roll a six-sided fair die. What is the probability of an ace (1 spot)?

A. 1/6 B. 1/2 C. 5/6

2. What is the probability of the *opposite* event? (*Not* rolling an ace.)

A. 1/6 B. 1/2 C. 5/6

9/3/18 15

Example: Coin toss

3. Toss a coin 3 times. What is the probability of getting *exactly* 1 head?

A. 1/2 B. 1/8 C. 3/8

4. Toss a coin 3 times. What is the probability of getting *at most* 1 head?

A. 1/2 B. 1/8 C. 3/8

9/3/18 16

Sampling with replacement

1 2 3 4 5 6

If we draw two tickets at random *with* replacement, the box stays the same from draw to draw, and the probabilities of drawing particular numbers don't change.

If we draw it *without* replacement, the box, and therefore the probabilities of the tickets within, will change.

"At random" means that all the tickets are equally likely.

Conditional probability

- Imagine that we are drawing a ticket from the box *without replacement.*
- What is the probability of getting 1 on the first draw?
    - 1/6
- If we draw the ticket marked "2" on the first draw, what is the probability of getting 1 on the second draw?
    - 1/5
- This probability is called a ***conditional probability***. It puts a condition on the first card, and *changes* the universe (the sample space) of the next outcome: the second card.

Multiplication rule

- Conditional probability written as **P(B|A)**, read as "the probability of the event B, given that the event A has occurred"
- Chance that two things will both happen is the chance that the first happens, *multiplied* by the chance that the second will happen *given* that the first has happened.
- Ex.: Draw a card at random, from a standard deck of 52 (4 suits, 13 cards of each suit, 2-10, J, Q, K, A).
    - P(King of hearts) =?
- Draw 2 cards one by one, without replacement.
    - P(1st card is K of hearts)=?
    - P(2nd card is K of hearts| 1st is not) = ?

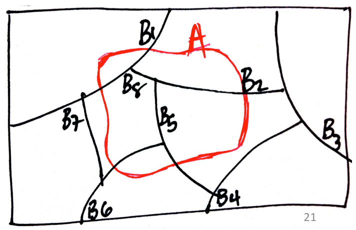
Rules governing chance

- If an event A is *certain* to occur, we say the chance of A (P(A)) is 100%.
- If an event A is *impossible*, we say P(A) =0%.
- The *opposite event* of A is called the *complement* of A and is denoted by A^C or "not A".
- $P(A^C)$ = P(not A) = 100% - P(A) = 1 – P(A)
- P(A) + $P(A^C)$ = 1 = 100%

Partitions

- We can always break our sample space into a bunch of mutually exclusive events (such that they don't overlap), that together make up the whole sample space.
- Such a collection of events is called a ***partition of S***, and any event A can be written as the union of its overlaps with the B_is.
)

$$A = (A\cap B_1)\cup(A\cap B_2)\cup\ldots(A\cap B_n)$$



Axioms of probability

- Now we can define the "axioms of probability":

1. The probability of any event is at least 0.
2. The probability of the sample space is 1.
3. If we have a partition of an event A given by A_1, A_2, …, A_n then the probability of A is the sum of the probabilities of the A_i.
4. Alternatively: If two events A_1 and A_2 are mutually exclusive, the chance that either A_1 happens or A_2 happens is the sum of the chances that each happens.

Probability distributions

- We talked last time about thinking about probability as a function on the subsets of S.
- We input a subset (including S and the empty set) and output a number in [0,1].
- Such a function is called a probability distribution over S.
- Why "distribution"? Notion of probability as a mass distributed over a discrete set or an area or volume.

Probability distribution example

- The **discrete uniform** probability distribution
- Let S contain n elements (where n > 0). For any subset A of S, define #(A) to be the number of elements of A. This is called the *cardinality* of A.
- Define P(A) = #(A)/n

- Examples of uniform distribution:
  - number of spots when a die is rolled,
  - number of heads in a single toss of a fair coin,
  - outcome of drawing a single ticket at random from a box of tickets.

Independence

- Multiplication rule:
  P(A and B) = P(A)P(B|A)

- Two things are *independent* if the chances for the second given the first are the same, no matter how the first turns out. That is, the chance of the second thing happening is *unaffected* by whether the first happened or not.
- If two events are not independent, they are called *dependent*.

- If A, B independent, then P(B|A) = P(B)

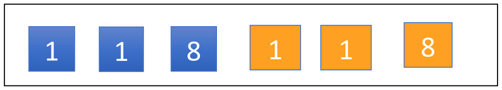
  P(A and B) = P(A)P(B)

Independence

- When sampling *with* replacement, draws are *independent*.

- When sampling *without* replacement, draws are dependent.

- Ex. (13.R.5): Draw 1 ticket at random. Are color and number independent? (Yes.)



Mutually Exclusive events

- Two events A and B are ***mutually exclusive*** when the occurrence of one precludes the other.
- To find the chance that at least one of two things will happen, check if they are mutually exclusive. If so, we can add the probabilities. If not, then what do we do?

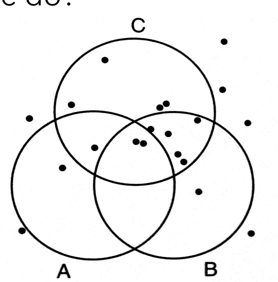


Figure from 14.B.2 (page 243) of text.

Two rules: Addition & Multiplication

- Addition rule: If *A* and *B* are mutually exclusive events, then the probability that at least one of the events will occur is the sum of their probabilities: $P(A \cup B) = P(A) + P(B)$
- If they are not mutually exclusive, does this still hold?
- How do we write the event that "at least one of the events *A* or *B* will occur? How do we draw it?
- $P(A \cup B) = P(A \text{ or } B) = \mathrm{P}(A) + P(B) - P(A \cap B)$

Two rules: Addition & Multiplication

- Multiplication rule: Chance that two events A and B will both happen is the chance that the first happens, *multiplied* by the chance that the second will happen *given* that the first has happened.
- Probability of A given B: $P(A \mid B)$
- $P(A \text{ and } B) = P(A \cap B) = P(AB)$
- $P(A \text{ and } B) = P(A \mid B) \times P(B)$

- If A and B are independent, then $P(A \mid B) = P(A)$
- Then we have that $P(A \text{ and } B) = P(A) \times P(B)$

9/4/18 29

Independent and Mutually Exclusive?

- If two events are mutually exclusive (disjoint), it means that these events don't overlap. If one event occurs, the other does not.
- If two events are independent , it means that the fact that one event occurring has **no bearing** on whether the other occurs. You do not get any information about the second event.
- Can two events be both mutually exclusive *and* independent?

9/4/18 30

Mutually Exclusive vs. Independent

- Make sure you understand the difference; these are *very* different ideas, though both apply to pairs of events.

- "Mutually exclusive" events means that the occurrence of one **prevents** the occurrence of the other. (This means that it reduces the chance of the other occurring to 0.)

- "Independent" events means that the occurrence of one **does not change** the chance of the other occurring.

- Read sections 14.3, 14.4 carefully.

9/4/18 31

Rules for computing probabilities

1. Counting equally likely outcomes: If all the outcomes in a sample space are equally likely, then $P(A) = \frac{\#\ of\ outcomes\ in\ A}{total\ \#\ of\ possible\ outcomes}$

2. $P(A) = 1 - P(not\ A) = 1 - P(A^C)$

3. Multiplication rule: $P(\text{A and } B) = P(A \mid B) \times P(B)$

4. Addition rule: $P(A\ or\ B) = P(A) + P(B)$