STAB52: An Introduction to Probability

Joshua Concon

University of Toronto Scarborough – Summer 2017

Pre-reqs are MATA37, which is Calculus for Mathematical Sciences II. Instructor is Dr. Mahinda Samarakoon. I highly recommend sitting at the front since he likes to teach with the board and can have a bit of trouble projecting his voice in large lecture halls. If you find any problems in these notes, feel free to contact me at conconjoshua@gmail.com.

Contents

1 Wednesday, May 3, 2017

1.1 What is probability?

He wanted us to think about what it really was. i.e. What does it mean for something to have a 50% probability? If a coin landing on heads has a 50% probability, it doesn't guarantee that such an event would occur if you flipped a coin twice.

After a few guesses, he gave us a definition.

1.2 Relative Frequency Definition of Probability

> **Relative Frequency**
>
> Consider the case where an experiment is performed n times.
>
> Let $|A|$ be the number of trials resulting in 'event' A.
>
> The **Relative Frequency** of $A = \frac{|A|}{n} = \gamma_n$

This is the Probability of A when n is large γ_n by itself is not an accurate definition of the probability of A occurring, so we take the limit as n goes to infinity and then we have such:

$$\lim_{n \to \infty} \gamma_n$$

This definition is difficult to use in most cases but relatively easy to understand. So here is a definition that is easier to do calculations with:

1.3 Formal Definition of Probability

He doesn't actually get into the definition of probability this lecture, instead he goes through a few terms that we need to define before we get to this definition.

> ### Sample Space (S)
>
> Sample Space (S): the set of all possible outcomes in an experiment.
> Size of S is denoted by $n(S)$, $\#S$ or $|S|$

<u>ex.</u> Tossing a coin once: $S = \{H, T\}$, $n(S) = 2$

<u>ex.</u> Rolling a 6-sided die: $S = \{1, 2, 3, 4, 5, 6\}$, $n(S) = 6$

<u>ex.</u> Tossing 2 different coins: $S = \{HH, HT, TH, TT\}$, $n(S) = 4$

> ### Events
>
> Subsets of a sample space

<u>ex.</u> Experiment rolling a die, $S = \{1, 2, 3, 4, 5, 6\}$
$A = \{1, 2\}$, $A \subseteq S$, so A is an event of S.
$B = \{5, 7\}$, $B \nsubseteq S$, so B is not an event of S.

You can also describe events with words.

$C =$ the result is an odd number $= \{1, 3, 5\}$, $C \subseteq S$, so C is an event of S.

An event can be the entire sample space and the null set.

$D = S \subseteq S$, so D is an event of S.
$E = \varnothing \subseteq S$, so E is an event of S.

<u>Operations on Events</u> (Unless specified, valid for all events A, B)

1. Where $S = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2\}$, $B = \{2, 4, 5\}$
   $A \cup B$: elements in A or B
   <u>ex.</u> $A \cup B = \{1, 2, 4, 5\}$

2. Where $S = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2\}$, $B = \{2, 4, 5\}$
   $A \cap B$: elements in A and B
   <u>ex.</u> $A \cap B = \{2\}$
   And for $C = \{5, 6\}$
   $A \cap C = \varnothing$

> Recall: So essentially all of the logic laws from CSCA67 hold for these sets as well.
>
> i.e.
> $$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
> $$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

3. $(A \cap B)^c = (A^c \cup B^c)$ (where A^c is the complement of A, which is the elements in S that is not in A)

   ex.
   Consider $A = \{1, 2\}$ and $S = \{1, 2, 3, 4, 5, 6\}$
   $A^c = S - A = \{3, 4, 5, 6\}$

4. $(A \cup B)^c = (A^c \cap B^c)$

5. $A \cup \varnothing = A$

6. $A \cap \varnothing = \varnothing$

Event Space

Is a set of all the possible events of S (i.e All the possible subsets of the set S). Denoted P

ex. The Event Space of the set $\{1, 2, 3\}$ is
$\{\{ 1 \},\{ 2 \},\{ 3 \},\{ 1,2 \},\{ 2,3 \},\{ 1,3 \},\{ 1,2,3 \}\}$

2　Friday, May 5, 2017

2.1　Formal Definition of Probability

> **Probability Measure Function (PMF)**
>
> A function $P : S \longmapsto [0,1]$ defined on sample space S is called a **Probability Measure Function**. It satisfies the following axioms:
>
> 1. $0 \geq P(A) \geq 1$
>
> 2. $P(S) = 1$ where S is a sample space
>
> 3. For disjoint events $A_1, ..., A_n$ (disjoint means where $A_i \cap A_k = \varnothing$ $\forall i, k \in \mathbb{N}\ 1 \geq i < k \geq n$)
>   then $P(A_1) \cup A_2 \cup ... \cup A_n = P(A_1) + P(A_2) + ... + P(A_n)$
>
> 4. $P(\varnothing) = 0$

2.2　Properties of Probability Measure Functions

Dr. Mahinda ends up adding **Result 1** to the definition of the Probability Measure Function (4) to stay consistent with the book, but also adds it here as well. It will be left again as a property to stay consistent with the lecture and the numbering.

Result 1:　$P(\varnothing) = 0$

Result 2:　If $A_1, ..., A_n$ is a finite collection of disjoint events, then

$$P(\bigcup_{i=1}^{n} A_i) = P(A_1 \cup A_2 \cup ... \cup A_n)$$

$$= \sum_{i=1}^{n} P(A_i)$$

Proof. (Proof of Result 2)

Define $A_i = \varnothing$, $i \geq n+1$

So since there is a finite amount of events, A_i where $i \geq n+1$ are just the null set

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$$

$$\bigcup_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} P(A_i)$$

By Axiom 3 of the PMF

$$= \sum_{i=n+1}^{\infty} P(A_i) + \sum_{i=1}^{n} P(A_i)$$

$$= \sum_{i=n+1}^{\infty} P(\varnothing) + \sum_{i=1}^{n} P(A_i)$$

$$= 0 + \sum_{i=1}^{n} P(A_i)$$

By Axiom 4 of the PMF

$$= \sum_{i=1}^{n} P(A_i)$$

∎

Result 3:   If $A_1, A_2, A_3, \ldots$ is a partition of S and B is any event then

$$P(B) = \sum_{i=1}^{\infty} P(A_i \cap B)$$

> ### Partitions
>
> A parition is a collection of events satisfying the following axioms:
>
> 1. $A_1, A_2, A_3, \ldots$ are disjoint events
>
> 2. $\displaystyle\bigcup_{i=1}^{\infty} A_i = S$

Proof. (Proof of Result 3)

For an arbitrary event B, Consider the partition $A_1, A_2, A_3, ...$ of S
The events $B \cap A_1, B \cap A_2, B \cap A_3, ...$ are, by the definition of partitions...

1. disjoint

2. $\bigcup\limits_{i=1}^{\infty} B \cap A_i = B \cap S = B$

Therefore
$$P(B) = P(\bigcup_{i=1}^{\infty} B \cap A_i) = \sum_{i=1}^{\infty} P(A_i \cap B)$$

By Axiom 3 of the PMF ■

Set Difference

For arbitrary sets A, B
$A \setminus B$ or $A \cap B^c$ refers to everything in A but not in B.

Result 4: $P(A \setminus B) = P(A) - P(A \cap B)$

Proof. (Proof of Result 4)

Note that $A \setminus B$ and $A \cap B$ are disjoint
$P((A \setminus B) \cup (A \cap B)) = P(A \setminus B) + P(A \cap B)$
Since $A \setminus B$ is disjoint from $A \cap B$
So for arbitrary events A, B
$P(A) = P(A \setminus B) + P(A \cap B) \implies P(A \setminus B) = P(A) - P(A \cap B)$ ■

<u>Note:</u> if $B \subseteq A$ then $P(A \setminus B) = P(A) - P(B)$

Proof. (Proof of Note)

if $B \subseteq A$ then $A \cap B = B$
therefore $P(A \setminus B) = P(A) - P(A \cap B) = P(A) - P(B)$ ■

Result 5:   If $A \subseteq B$ then $P(A) \leq P(B)$

Proof. (Proof of Result 5)

$$A \subseteq B \Longrightarrow P(A \setminus B)$$
$$= P(B) - P(A) \geq 0$$

By Axiom 1 of the PMF

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

∎

Result 6: $P(A^c) = 1 - P(A)$

Proof. (Proof of Result 6)
Consider an arbitrary event $A \subseteq S$
Since we know that if $A \subseteq S$ then $P(A \setminus B) = P(B) - P(A)$
and that $A^c = S \setminus A$

$$P(A^c) = P(S \setminus A) = P(S) - P(A)$$
$$= 1 - P(A)$$

∎

Result 7:   we know $P(A \cup B) = P(A) + P(B)$ if A, B are disjoint and $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ for any 2 events A, B.

Proof. (Proof for Result 7)
since we know that $A \cup B = A \cup (B \setminus A)$ and that events A and $(B \setminus A)$ are disjoint.

$$P(A \cup B) = P(A \cup (B \setminus A))$$
$$= P(A) + P(B \setminus A)$$
$$= P(A) + P(B) - P(A \cup B)$$

By Result 4

∎

<u>Note:</u> This result can be extended to more than 2 events <u>e.g.</u>

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Example 1: Consider a bag with 3 red marbles and 7 green marbles inside the bag. What is the probability that you would select a red marble from the bag at random? How about the probability of selecting a green marble?

Solution 1: Simply consider all possible events, and then for each probability x, take all the events where x happens and divide it by the sample space. Let R, G represent the events where a red and green marble respectively is selected.

$$P(R) = \frac{R}{R+G} = 0.3$$

$$P(G) = \frac{G}{R+G} = 0.7$$

Example 2: Consider a class of 33 students, 17 of them aced the midterm, 14 of them aced the final and 11 did not have As on either. How many of them aced both the midterm and the final?

Solution 2: Let A, B refer to the students who have aced the midterm and final respectively. We are trying to solve for $P(A \cap B)$.

$$P(A) = 17/33$$
$$P(B) = 14/33$$
$$P(A \cup B) = 22/33$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$22/33 = 17/33 + 14/33 - P(A \cap B)$$
$$22/33 = 31/33 - P(A \cap B)$$
$$22/33 + P(A \cap B) = 31/33$$
$$P(A \cap B) = 31/33 - 22/33$$
$$P(A \cap B) = 9/33$$

2.3   Uniform Probability Measure on Finite Sample Spaces

> **Uniform Probability Measure on Finite Sample Spaces (UPMFSS)**
>
> If S is a finite sample space (i.e. $|S| < \infty$) of equally likely outcomes, and A is any event, then $P(A) = \frac{|A|}{|S|}$

But is this sample space a probability measure?

Proof. (Proof of UPMFSS being a probability measure)

1. $\underline{0 \leq P(A) \leq 1}$ for an arbitrary event A

$$\varnothing \subseteq A \subseteq S$$
$$|\varnothing| \leq |A| \leq |S|$$
$$0 \leq \frac{|A|}{|S|} \leq 1$$

2. $\underline{P(S) = 1}$

$$P(S) = \frac{|S|}{|S|} = 1$$

3. $\underline{P(A_1) \cup A_2 \cup ... \cup A_n = P(A_1) + P(A_2) + ... + P(A_n)}$
   Let $A_1, A_2, ...$ be disjoint events
   Consider $|\bigcup\limits_{i=1}^{\infty} A_i|$

$$|\bigcup_{i=1}^{\infty} A_i| = |A_1| + |A_2| + ...$$

   Since the A_i's are disjoint

$$\frac{|\bigcup\limits_{i=1}^{\infty} A_i|}{|S|} = \frac{|A_1|}{|S|} + \frac{|A_2|}{|S|} + ...$$
$$P(|\bigcup_{i=1}^{\infty} A_i|) = P(A_1) + P(A_2) + ...$$

∎

2.4   Counting (Combinatorics)

Example 1: How many 4 digit numbers can be formed from digits $1, 2, ..., 9$ if repeats are not allowed?

Solution 1: $P(9,4) = \frac{9!}{4!} = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5$ since there are 9 numbers to choose from and there are 4 digits that these numbers can "occupy".

Example 2: Find the probability that the 4-digit number formed from the question above is divisible by 5.

Solution 2: $\frac{P(8,3)}{P(9,4)} = \frac{8 \cdot 7 \cdot 6}{9 \cdot 8 \cdot 7 \cdot 6} = \frac{1}{9}$ since there are $P(9,4)$ combinations in total and the numbers divisible by 5 end with a 5 (since it is not possible for the number in the question above to end in a 0). So there are 3 digits left to be assigned a number with only 8 numbers since 5 is already used, so the numerator is $P(8,3)$.

Counting Subsets: The number of subsets of size k that can be formed from n objects is

$$\frac{n!}{k!(n-k)!} = C(n,k) = \binom{n}{k}$$

3   Wednesday, May 10, 2017

3.1   Counting (continued)

Example 1: A professor will select 5 out of 10 questions for an exam, and a student studies the answers to 6 of these 10 questions. What is the probability that all the questions that the professor will select questions are questions that the student has studied for?

Solution 1: The sample space $|S| = \binom{10}{5}$ and the cases that all the questions that the professor will select questions that the student has studied for is $\binom{6}{5}$ so the probability is:

$$\frac{\binom{6}{5}}{\binom{10}{5}}$$

Example 2: There are 20 members in a group. We want to divide this group into 3 sub-groups (say, A, B, C) of sizes $8, 8, 4$ respectively. In how many ways can we do this?

Solution 2: We choose 8 for A, and then 8 from the remaining 12 for B where the order does not matter and the rest go to C, so we have.

$$\binom{20}{8}\binom{12}{8}\binom{4}{4} = \frac{20!}{8!12!} \cdot \frac{12!}{8!4!} \cdot \frac{4!}{4!} = \frac{12!}{8!8!4!} = \binom{12}{8\ 8\ 4}$$

Example 3: A deck of 52 cards has 12 picture cards. We want to divide this among 4 players (say, A, B, C, D). What is the probability that each player will get exactly 3 picture cards?

Solution 3: The sample space $|S| = \binom{52}{13\ 13\ 13\ 13}$ The cases for distributing all of the 12 picture cards between the 4 players is $\binom{12}{3\ 3\ 3\ 3}$ and the cases for distributing all of the 40 regular cards between the players is $\binom{40}{10\ 10\ 10\ 10}$ so the total probability is:

$$\frac{\binom{12}{3\ 3\ 3\ 3} \cdot \binom{40}{10\ 10\ 10\ 10}}{\binom{52}{13\ 13\ 13\ 13}}$$

3.2   Conditional Probability

> **Conditional Probability**
>
> Let A, B be two events such that $P(B) > 0$, then the conditional probability of A given B is:
>
> $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Proof. (Proof that $P(\cdot|B)$ is a probability measure)
Axiom 1

$$\varnothing \subseteq A \cap B \subseteq B$$
$$P(\varnothing) \leq P(A \cap B) \leq P(B)$$
$$0 \leq \frac{P(A \cap B)}{P(B)} \leq 1$$

Therefore $P(\cdot|B)$ satisfies axiom 1.

Axiom 2
Let S be the sample space, and let B be any event

$$S \cap B = B$$
$$P(S \cap B) = P(B)$$
$$P(S|B) = \frac{P(S \cap B)}{P(B)} = 1$$

Therefore $P(\cdot|B)$ satisfies axiom 2.

Axiom 3
Let $A_1, A_2, ...$ be a collection of disjoint sets

Note that $\forall$ events B that $A_1 \cap B, A_2 \cap B, A_3 \cap B, ...$ are also disjoint.

$$P((A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup ...) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + ...$$
$$= P((A_1 \cup A_2 \cup A_3 \cup ...) \cap B)$$
$$\frac{P((A_1 \cup A_2 \cup A_3 \cup ...) \cap B)}{P(B)} = \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} + \frac{P(A_3 \cap B)}{P(B)} + ...$$
$$P(A_1 \cup A_2 \cup A_3 \cup ...|B) = P(A_1|B) + P(A_2|B) + P(A_3|B) + ...$$

Therefore $P(\cdot|B)$ satisfies axiom 3.
Therefore $P(\cdot|B)$ is a probability measure. ■

Example 4: Consider the two events A, B such that
$P(A) = 0.5, P(B) = 0.3, P(AB) = 0.1.$ Find...

1. $P(A|B) = \frac{P(AB)}{P(B)} = \frac{0.1}{0.3} = \frac{1}{3}$

2. $P(B|A) = \frac{P(AB)}{P(A)} = \frac{0.1}{0.5} = \frac{1}{5}$

3. $P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A)+P(B)-P(AB)} = \frac{0.5}{0.5+0.3-0.1} = \frac{5}{7}$

4. $P(A|A \cap B) = \frac{P(A \cap A \cap B)}{P(A \cap B)} = \frac{P(A \cap B)}{P(A \cap B)} = 1$

5. $P(AB|A \cup B) = \frac{P(AB \cap (A \cup B))}{P(A \cup B)} = \frac{P(AB)}{P(A \cup B)} = \frac{0.1}{0.5+0.3-0.1} = \frac{1}{7}$