STOR 435.001 Lecture 4

Axioms of Probability - II

Jan Hannig

UNC Chapel Hill

How can we introduce and think of probabilities of events?

Natural to think: repeat the experiment n times under same conditions; let $n(E)$ be the number of times event E occurs in these n repetitions; think

$$\lim_{n\to\infty} \underbrace{\frac{n(E)}{n}}_{\text{frequency of } E \text{ in } n \text{ rep.}} = \underbrace{\mathbb{P}(E)}_{\text{probability of } E} \qquad (*)$$

For us $(*)$ is not a definition.

Why not simply say probability IS long run proportion?

- **Example:** Why not just say that “Probability of a coin turning up heads” means long run proportion of heads is 50%?
- Problem with this definition: How long is long? Is 100 times enough? Or 1000? So as a precise definition this does not make sense.
- Second issue: This long run frequency type definition cannot allow us to model questions like *“What is the chance stock price of Amazon rises by at least 10% next year”* or *“what is the chance Brad has an accident whilst driving next year?”*. “Keep repeating the experiment makes no sense in such settings”
- Subjective probability can be realized by bets and prices, e.g., the premium Brad pays.

Punchline in quantitative fields

- We *first* build a model of a real world system. (Set up events and their probabilities.)
- If the model accurately reflect reality **and** if we are able to repeat the experiment many times *then* the long run proportion will match the probability.
- **Name of the game:** Build accurate models of reality (and train it on good and plentiful data).
- In examples such as dice etc there is a consensus on the *correct* model of reality. In other “real world” situations much more complicated. The better modeler will get the Moolah.
- Example: *The Netflix prize*

Roadmap

Modern axiomatic approach to Probability Theory

- Start with simpler, more evident assumptions for probabilities
- Then show $(*)$ (namely in the long run proportion of times an event happens converges to the probability of the event).
- Sir David Spiegelhalter `@d_spiegel` urges communication of probabilities through expected frequencies
- Probability of 0.2 expressed as out of 100 we expect 20.

Steps

Steps

1. We start by defining a “sample space” S namely collection of **all** possible outcomes of the experiment.
2. Then specify collection of subsets $\mathcal{F}$. Each of these will be called events. We will be “only allowed” to talk about probability of these events. When collection of outcomes finite or countable just take $\mathcal{F}$ to be the power set (all possible subsets). For the rest of the course don’t worry about this. Just keep in mind that you are intersted in a collection of events and calculating probabilities for these events.
3. Math model: For each event in $\mathcal{F}$ will assign a “probability” of the event. This assignment has to satisfy three natural axioms.

Axioms of Probability

Let S be a sample space, and $\mathbb{P}(E)$ denote the probability of event E.

Assign $\mathbb{P}(E)$'s so that the following axioms hold:[1]

Axiom 1: $0 \le \mathbb{P}(E) \le 1$ for any event E.

Axiom 2: $\mathbb{P}(S) = 1$.

Axiom 3: for mutually exclusive (disjoint) events $E_1, E_2, \ldots$:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Note: Axioms? In principle, they need to be verified when probabilities are assigned.

[1] Andrey Kolmogorov (1903-1987), Russian mathematician, 1930's.

Axioms of Probability

Consequence 1:

$\mathbb{P}(\emptyset) = 0$.

☼

Consequence 2:

If $E_1, E_2, \ldots, E_n$ are mutually exclusive (disjoint), then

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathbb{P}(E_i).$$

In particular, for mutually exclusive A and B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Note: this makes sense since for such A and B, $n(A \cup B) = n(A) + n(B)$.

☼

Example 1:

$S = \{H, T\}$. Possible events: $\{H\}$, $\{T\}$, S, $\emptyset$. Fix $p \in (0, 1)$ and assign $\mathbb{P}(\{H\}) = p$, $\mathbb{P}(\{T\}) = 1 - p$, $\mathbb{P}(S) = 1$, $\mathbb{P}(\emptyset) = 0$. Check the axioms.

☼

Note: If $\mathbb{P}(\{T\}) = 1 - \frac{p}{2}$, which of the axioms are not satisfied?

☼

Example 2:

(i) $S = \{x_1, x_2, \ldots, x_N\}$.
E.g. $S = \{1, 2, 3, 4, 5, 6\} = \{$outcomes in rolling a die$\}$.

(ii) Assign $\mathbb{P}(\{x_i\}) = p_i$, $p_i \in (0, 1)$ with $p_1 + p_2 + \ldots + p_N = 1$.
E.g. $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 1/6$ above.

(iii) E.g. Axiom 3 implies that
$\mathbb{P}(\{x_2, x_4, x_6\}) = \mathbb{P}(\{x_2\}) + \mathbb{P}(\{x_4\}) + \mathbb{P}(\{x_6\}) = p_2 + p_4 + p_6$.

(iv) For any event E, assign

$$\mathbb{P}(E) = \sum_{i:x_i \in E} \mathbb{P}(\{x_i\}) = \sum_{i:x_i \in E} p_i.$$

With this assignment, Axioms 1 and 2 are satisfied. One can show that Axiom 3 is satisfied as well.

Example 3

You have an unfair dice such that

$$\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = 1/4$$

and

$$\mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = 1/12$$

What is the chance that when you throw the dice, you get an even number?

☼

Special case of Example 2
Probability model where all outcomes equally likely ⚠

(i) Sample space $S = \{x_1, x_2, \ldots, x_N\}$.

(ii) $p_1 = p_2 = \ldots = p_N = \frac{1}{N}$ (equally likely outcomes), in which case

$$\begin{aligned}\mathbb{P}(E) &= \frac{\text{Number of outcomes } x_i \text{ in } E}{\text{Total number of outcomes } x_i \text{ in } S} \\ &= \frac{\text{Number of outcomes } x_i \text{ in } E}{N}.\end{aligned}$$

(iii) Note: We will often use combinatorics to compute the numbers in the numerator and denominator.

Example 4 (a)

In a small town with 900 adults, there are 600 Democrats (say $D_1, D_2, \ldots, D_{600}$) and 300 Republicans (say $R_1, R_2, \ldots R_{300}$). You want to sample 2 people from this population. To do this you fill a box with the names of the people and then pick 2 pick names at random without replacement. What is the probability model?

Example 4 (b)

Pick 2 names at random with replacement. What is the probability model?

Axioms of Probability

Terminology: $(S, \text{events}, \mathbb{P})$ or $(S, \mathbb{P})$ (with $\mathbb{P}$ satisfying Axioms 1,2 and 3) is called a **probability space** (**probability model**).

Note: Axioms will often be taken as rules, especially for equally likely outcomes.

Note: In the context of Venn diagrams, one can think of probability as area (or mass).

Next: Some simple propositions **valid for any probability model.**

Proposition 1:

$\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.

☼

Proposition 2:

If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$.

☼

Axioms of Probability

Proposition 3

For any two events E, F,

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F).$$

☼

Think-Pair-Share (with made up numbers)

Suppose 80% of UNC students have either a Facebook account or a Twitter account. Suppose you know 60% have a Facebook account and 25% have a Twitter account. If one of the students is selected uniformly at random, what is the chance they are on both Facebook and Twitter?

A 80%

B 60%

C 25%

D 5%

Axioms of Probability

More than 2 events

e.g. with 3 events,

$$\mathbb{P}(E \cup F \cup G) =$$

Note: There is an analogous formula for an arbitrary number of events, called **inclusion-exclusion identity**. See p. 31 in the textbook. For the course you will need to know the formula only upto and including 3 events.

Example

In Fall 2014, the department has a cohort of 100 MDS juniors. Students in the class are enrolled in various classes including STOR 435 (undergraduate probability), Math 547 (Linear Algebra for applications) and COMP 401 (Foundations of Programming). To simplify notation below I will denote these as Prob., LA and COMP. Assume the classes are not taught at the same time so there can be students who are taking all 3 courses. Further there can also be students who are taking none of these 3 courses. 65 of these students are in Prob, 45 in LA and 40 in COMP. There are 25 students that are in both Prob. and COMP, 20 that are in both LA and COMP, and 30 that are in both Prob. and LA. In addition, there are 10 students taking all 3 classes.
(a) If a student is chosen randomly, what is the probability that he or she is not in any of the three courses?
(b) If a student is chosen randomly, what is the probability that he or she is taking exactly one of the above three courses?

☼

Solution continued:

Example

(a) Your friend is hosting a birthday party with a 120 people. Being rich and generous, he is going to be giving 6 ipads away. To be fair (and not piss anyone off), he puts the names of all the guests in a big bowl and pulls out 6 names. What is the chance that you get selected?

(b) Do the same problem when you have n guests and k free ipads.

☼

Birthday problem

Our class has around 95 students. What is the probability that there are at least two people with the same birthday? You can assume that all the configurations are equally likely and you may assume no one in the class was born on February 29th.

Deck of cards

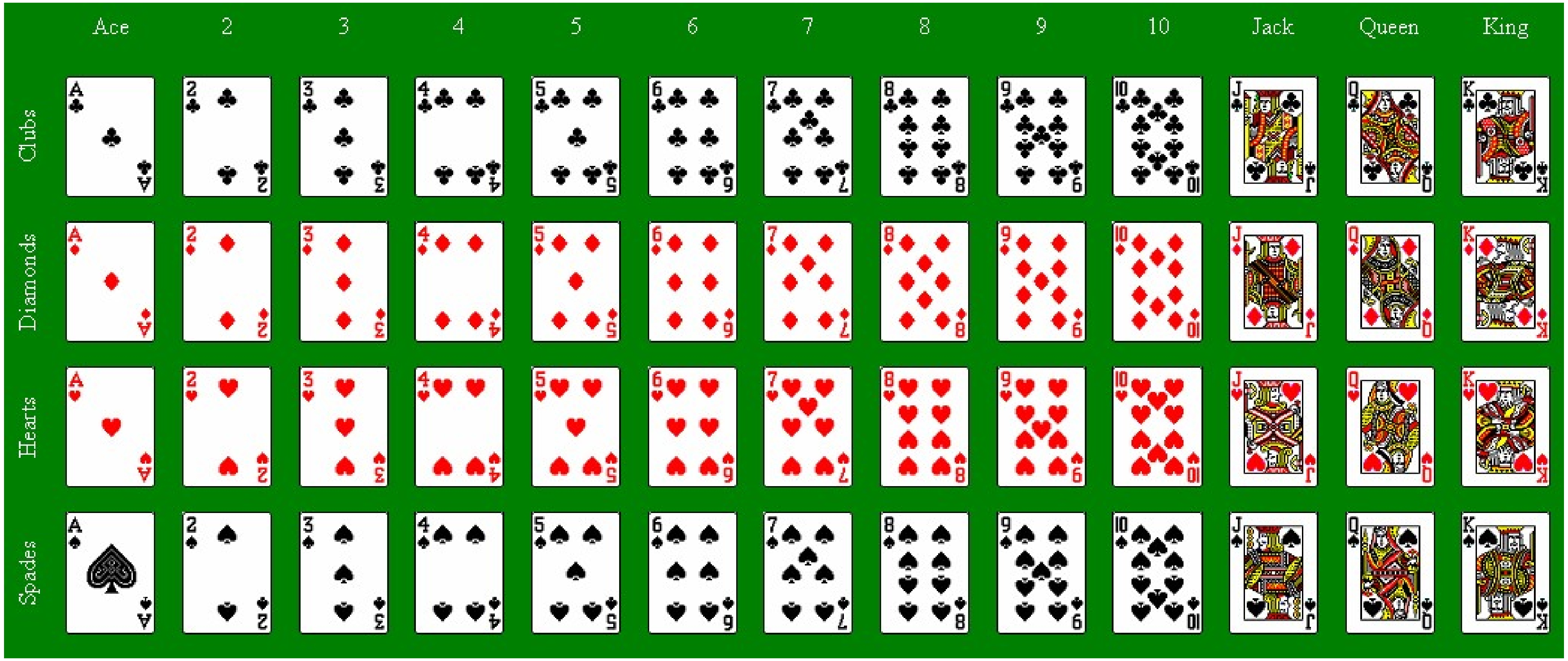

Figure: From davidbridge.com

Poker

A deck of cards consists of 52 cards. Each card has a denomination (Ace, 2,3,. . ., 10 Jack Queen King, so that there are 13 denominations) and 4 shapes ($\clubsuit$, $\diamondsuit$, $\heartsuit$, $\spadesuit$). While playing poker, you are given 5 cards from the deck of 52 cards. You are said to have a full house if you have 3 cards of the same denomination and 2 other cards of the same denomination, for example 3 Aces and 2 10's. What is the probability of you getting a full house?