Boston University
Questrom School of Business

MF 793 – Fall 2021

Eric Jacquier

QUICK OVERVIEW OF PROBABILITIES

We will go very fast on the first part of this note as you saw these during your prep. course

- **Great Book:** "Statistical Inference" by Casella and Berger
- **Key concepts**: Probability space, probability, conditional probability, independence
- Fundamental Result: Total Probability Theorem, Bayes Theorem

1. Two Types of Probability

- **Classical probability**

> If a random experiment can result in **n** mutually exclusive and equally likely outcomes, and
> If n_A of these outcomes have a characteristic A,
> The probability of A, p(A), is the fraction n_A / n

- Throw 2 dice
  Mutually exclusive equally likely outcome: Set of the two numbers on the two dice
  Characteristic A could be: "The sum is larger than four"

- Throw a coin twice:
  Three mutually exclusive outcomes: 2 Heads, a H and a Tail, 2 Tails
  Say A is: We get two heads p(A) = 1/3 ?

  Problem 1: What if the outcomes are **not** equally likely?

- A: a positive integer is even p(A) = 1/2 ?

  Problem 2: What if the number of possible outcomes is **infinite**?

What is the probability that a cookie has more than three raisins?

We need to extend the classical probability to allow for **non equally likely** individual outcomes

- **Frequentist probability**

Run the experiment of throwing the coin many times

Assume that there exists a number p which is the probability of Heads

Approximate p as the fraction of Heads in the large experiment

Complication:

Need to be able to run the experiment **many times** in pretty stable conditions

Problem:

What if we are interested in a future outcome for which there is no or little data?

Eg. The Russians lost a nuclear submarine in the Pacific, what is the probability it is within 50 miles of Honolulu. Yikes!

- **Subjective probability**

Will free us from this problem … To be seen in MF840: Bayesian Statistics

2. Probability Models

Let's think about this conceptual experiment ..

Every conceivable *event* .... an event is a sample point ω

The totality of the conceivable events is the sample space Ω

We may be interested in *(many different)* subsets of the sample space,

- **Connection to Set theory**

Event A **or** Event B: One **or** the other **or maybe both** happen Union $A \cup B$

Event A **and** Event B: both happen Intersection $A \cap B$

Not A: Complement $\bar{A}$

The collection of all possible subsets of Ω is an Algebra $\mathcal{A}$

- A **Probability** is

  a set function with domain $\mathcal{A}$ and values in [0,1] satisfying

  $P(A) \geq 0, \quad \forall A \in \mathcal{A}$

  $P(\Omega) = 1$

  $P(\cup A_i) = \sum P(A_i)$ for mutually exclusive A_is

- Probability Space: the triplet $(\Omega, \mathcal{A}, P)$

... OK for the fancy words, probability is mostly about counting !

For MF793, we will concentrate on counting!

3. Let us count

Jar with **M** distinguishable balls – marked 1 to M for example, or with different colors

- # ordered samples of **n** balls drawn **with** replacement: M^n [1]
- # **ordered samples** of **n** balls drawn **without** replacement: ***M (M-1) (M-2) x ... x (M-n+1)***

  That is: $\frac{M!}{(M-n)!}$ [2]

- # possible ways to **order** n objects? *n (n-1) (n-2) x … x 1 = n!* [3]
- # subsets of size n in the jar ?

  We have $\frac{M!}{(M-n)!}$ ordered samples of size n

  A subset is **not** an ordered sample: Balls order does not matter: (1,2) is the same as (2,1)

  So there are $\frac{M!}{(M-n)!n!}$ **subsets** of size n in the urn, aka **number of combinations**

This is the coefficient of $\mathbf{a^n\, b^{M-n}}$ in the expansion of $(a+b)^M$: $\binom{M}{n}$ because it is the number of times that n a's come up in the polynomial.

- Prove that the sum of the binomial coefficients of $(a+b)^M$ is 2^M
  - How many **subsets** does an urn of M balls contain

    Number of subsets: **2^M** Why ?

    Each of the M balls can be in the subset or not: 2 choices for each ball

    But this includes the empty set (no ball at all):

    Excluding the empty set, we have **2^M-1** subsets of balls in the urn,

  - So, 2^M is the sum of the binomial coefficients !

    The number of subsets of 0, 1, 2, ..., n, ... , M-1, M, balls in the urn

    .

4. Multiple Events, Fundamental Definitions

- **Conditional Probability of A given B**, written P(A | B) is defined as

$$P(A|B) = \frac{P(AB)}{P(B)} \qquad if\ P(B) > 0 \qquad \textbf{[4]}$$

A and B are two events in Ω

P(AB): joint probability of A and B both occurring

P(A|B): conditional probability of A occurring given that B has occurred

Definition compatible with frequency probability definition (counting): $P(A|B) = N_{AB} / N_B$

Question: Toss two coins A and B,

P (2 Heads | H_A) = ½ obviously = (by definition) /

P(2 Heads | at least 1 Head) = /

You need to be able to go from joint probabilities to conditional probabilities … and back!

- **Total Probability Theorem**

  B_i collection of non-empty, mutually disjoint events satisfying $\cup B_i = \Omega$

$$\forall \boldsymbol{A} \in \boldsymbol{\mathcal{A}},\ \boldsymbol{P(A) = P(\cup AB_i) = \sum P(AB_i) = \sum P(A|\, B_i)\, P(B_i)} \qquad \textbf{[5]}$$

  Obvious with sets, the idea: breaking A along all its intersection with the B_is

  Corollary:
$$P(A) = P(A|B)\, P(B) + P(A|\bar{B})\, P(\bar{B}) \qquad [5']$$

- **Multiplication Rule:**

$$\mathbf{P(A_1\, A_2\, \ldots\, A_n) = P\,(A_n|\, A_{n-1}, \ldots\, A_1)\, P(A_{n-1}|A_{n-2},\ldots,A_1)\, \ldots\, P(A_2|A_1)P(A_1)} \qquad \textbf{[6]}$$

  Generalizes the conditional probability: p(A B) = p(A|B) p(B)

  Used in econometrics, we will think of the As as the densities of data points.

- **Independence: A and B are independent if P(A|B) = P(A)**

  P(B) is the same whether or not we condition on A

  Same as .. if **P(A) P(B) = P(AB) or P(B|A) = P(B)** **[7]**

5. Bayes Theorem

$$\boldsymbol{P(B|A) = \frac{P(A|B)\,P(B)}{P(A)}} \qquad [8]$$

- Foundation of modern econometrics
- Crucial for two-stage experiments and medical statistics
- We often have A|B but we want the answer to B|A. Bayes Theorem gets the answer.
- Many reporting research confuse the two conditional probabilities (A|B and B|A) !
- Bayes Theorem is used to get the (interesting) reverse probability

Doctors see many lung cancer patients, ask them if they smoke, create a data base.

They estimate P(smoking | cancer)

But they want to know: P(cancer | smoking) and P(cancer | non-smoking)

- Example: Medical testing, we have a test to detect Tuberculosis:

  p(TB) = 0.001. The unconditional probability observed in the entire population

  Test Accuracy: test detects TB with probability 0.999 if subject has TB

  Test False positive: test is positive with probability 0.002 in healthy subject

  Test reacted positive on Johnny, what is the probability Johnny has TB?

  P(Johnny has TB | positive test) =

  Test reacted negative on Mary,

  P(Mary has TB | negative test) =

- Note how using Bayes Theorem most always requires an application of the Total Probability Theorem.

Questions

- Probability that at least 2 students in a class of 60 have the same birthday ?
- Is the above question simpler than "...exactly 2 students..." ?
- **Famous Envelope puzzle**:

  A swami randomly puts m dollars and 2m dollars in two envelopes
  You don't know m.
  You and your friend get one envelope.
  You open your envelope and find x dollars (you know x, since you found it!)

  Should you get your friend's unopened envelope or keep your x dollars?

  Answer: what is your expectation of $y, the amount in your friend's envelope

  Warning: There is no simple mathematical answer without an assumption

- **St Petersburg Game Paradox**: How much would you pay to enter the following game?

  A fair coin is tossed sequentially until...

  ... the first head appears, after N Tails, you earn 2^N Rubles and the game ends.

  Hint: What is the expected earning of this game ?