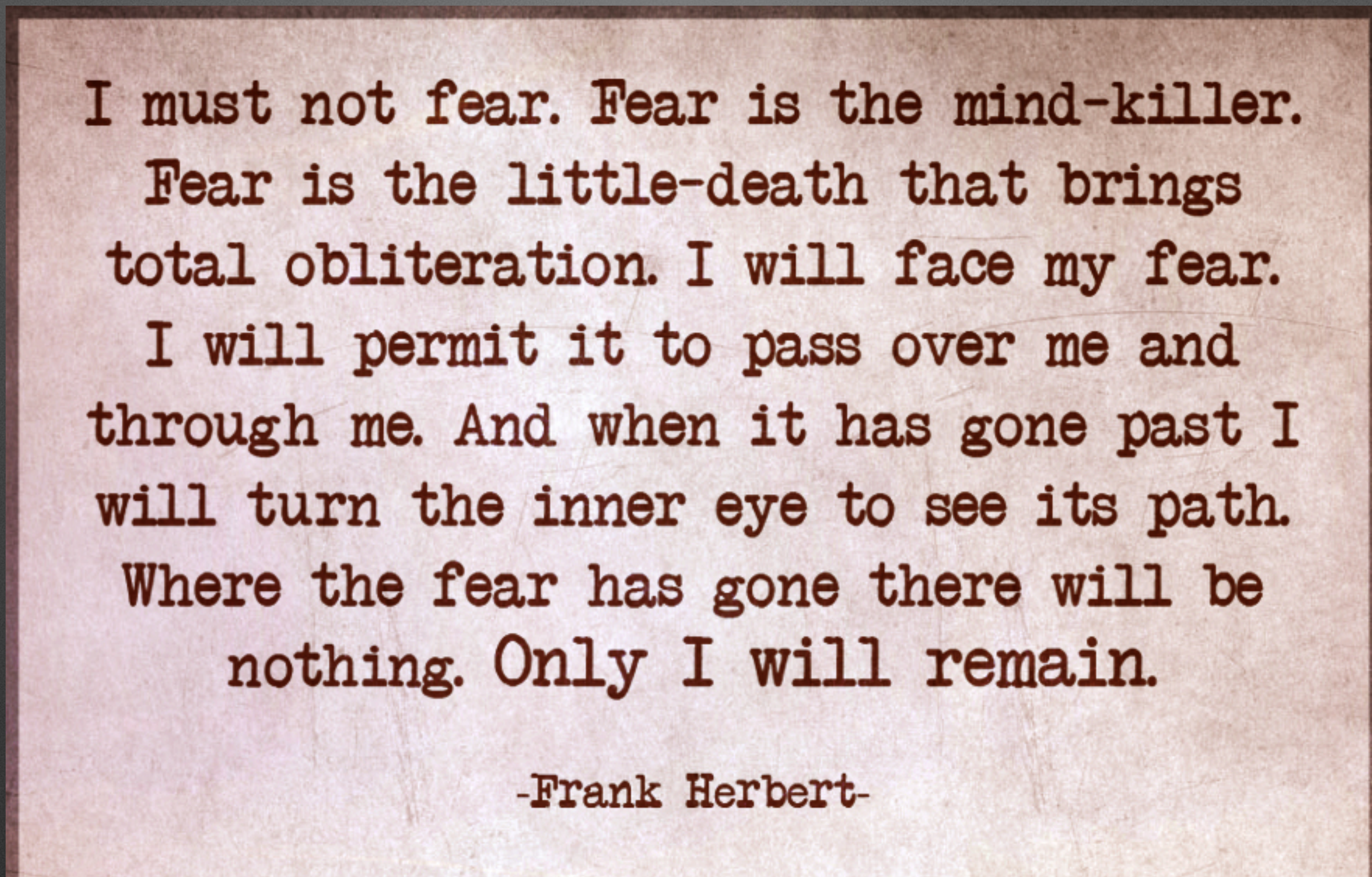
A Fear Based Approach to Probability

Joe

Introduction

The Mind-Killer



Session Objective

Refresh: https://github.com/gSchool/DSI_Lectures
we'll be using probability/jGartner. Needless to say, I'm taking a bit of a different approach to todays lectures… I took some stuff from Miles Erickson's feel free to look

1. Define fundamental laws of:
   set notation
   combinatorics
   probability

2. Use these laws to solve problems!

Terminology

Fear inducing problem 1 - define probability without using probability or likelihood :-)

Do the same for statistics.

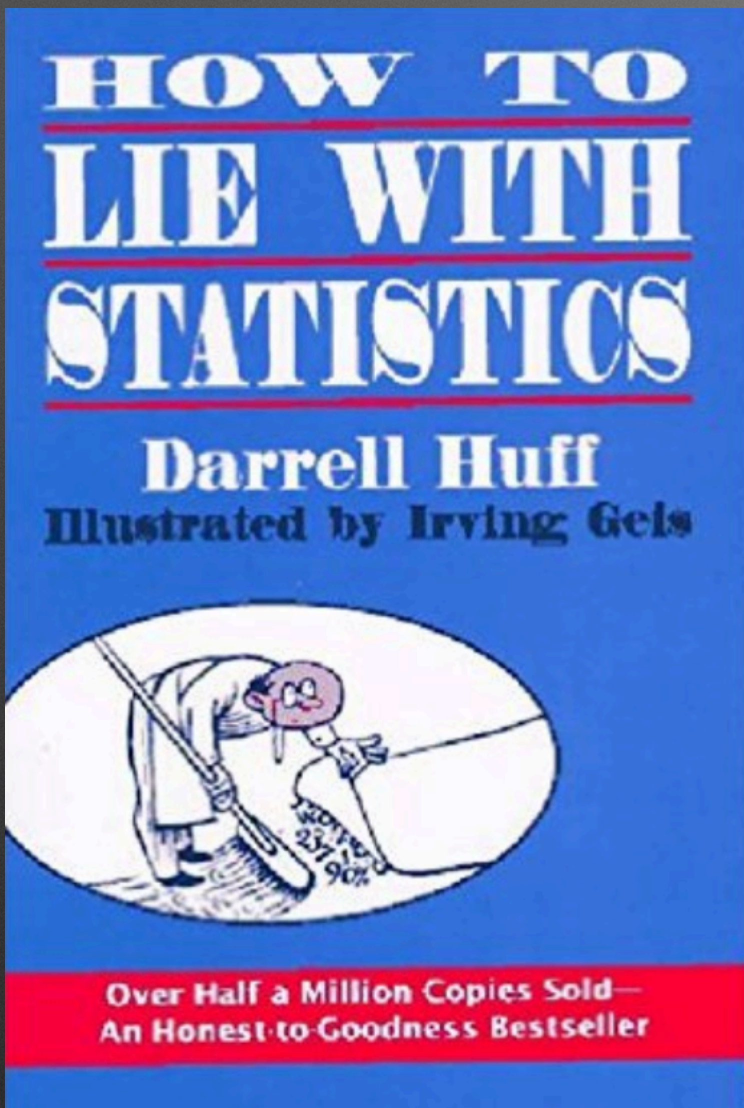
Terminology

Probability can be defined as a degree of belief that some random event will occur in the future

Likelihood is the posterior evaluation of probability (e.g. “What is the likelihood of you being accepting to this program?”)

Statistics are a set of techniques used to infer the true characteristic of a group by the measurement of a portion of the group

Motivation: Why do I need to Know Probability



- Important for exploratory data analysis
  - E.g. I want to build a model to predict user churn in twitter. How do I know if a user has left?
- Most ML has some basis in statistics.

Set Notation

Set Questions

- Your friend has two children and you know one is a girl. If there are equal number of men and women in the world, what is the probability both are girls?
- If the probability that a person drinks coffee is .7 and the probability that a person drinks teas is .4, can these sets be disjoint?
- Suppose that Rick is not (interested and impressed) by your attempts at answering those questions. How many quantum ricks did we just make?

Lingua Franca

Some definitions

- A set S consists of all possible outcomes or events and is called the sample space
- Union: $A \cup B = \{x : x \in A \sim or \sim x \in B\}$
- Intersection: $A \cap B = \{x : x \in A \sim and \sim x \in B\}$
- Complement: $A^{C} = \{x : x \notin A\}$
- Disjoint: $A \cap B = \varnothing$
- Partition: a set of pairwise disjoint sets, A_j, such that $\cup^{\infty}{}_{j=1} A_j = S$
- DeMorgan's laws: $(A \cup B)^{C} = A^{C} \cap B^{C}$ and $(A \cap B)^{C} = A^{C} \cup B^{C}$

Let’s do some white boarding!

Combinatorics

Not So Scary Questions

- Suppose we invent a simple lottery, where we have 20 possible numbers. What is the probability to win if:
  - You have to select a single number?
  - You have to select three numbers?
  - You have to select three numbers <u>and</u> the order they will be selected in (i.e. 1,2,3 != 2,1,3) ?

Permutation

Pronounced 'n factorial', this describes the number of ways to order n <u>unique</u> objects.

$$\begin{aligned} n! &= \prod_{k=1}^{n} k \\ &= 1 \cdot 2 \cdot 3 \cdots (n-2) \cdot (n-1) \cdot n \\ &= n(n-1)(n-2) \cdots (2)(1) \end{aligned}$$

Combinatorics

pronounced
"n choose k" describes the number of combinations of size k that can be made from a sample of n elements

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

Combinatorics

How many ‘ordered’ combinations can I make

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

Combinatorics

Remove ‘redundant’ instances that are just perutations of one another. If order matters, we simply remove this term.

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

Data Science Usage

- How many unique solutions to x+y+z=5 exists, assuming x, y, and z are non negative integers?
  Case - L1 Regularization

- For some clustering problem there are 5 ‘true’ groups, but I run the algorithm trying to find 7. If I rerun the algorithm with different random seeds, how many unique combination can I get?

- Is it possible to have a statistically significant test with only 8 trials? Yes - Fisher exact test
  - On your own - look up the “lady drinking tea” problem

Law’s of Probability

Probability Questions

- What’s the probability of two dice adding to 10 or higher, given that at least one reads 6?
- What’s the probability one having at least one 6, given the sum of dice is 10 or higher?

More Definitions

Definition of probability

Given a sample space S, a *probability function* P of a set has three properties.

- $P(A) \geq 0 \; \forall \; A \subset S$
- $P(S) = 1$
- For a set of pairwise disjoint sets $\{A_j\}$, $P(\cup_j A_j) = \sum_j P(A_j)$

Independence

Two events A and B are said to be *independent* iff

$$P(A \cap B) = P(A)P(B)$$

or equivalently

$$P(B \mid A) = P(B)$$

so knowlege of A provides no information about B. This can also be written as $A \perp B$.

Total Probability

Law of Total Probability

If B_n is a partition of all possible options, then

$$\begin{aligned} P(A) &= \sum_j P(A \cap B_j) \\ &= \sum_j P(A \mid B_j) \cdot P(B_j) \end{aligned}$$

Bayes' Theorem

Bayes' theorem

Bayes' therem says that

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Where A and B are two possible events.

To prove it, consider that

$$\begin{aligned} P(A \mid B)P(B) &= P(A \cap B) \\ &= P(B \cap A) \\ &= P(B \mid A)P(A) \end{aligned}$$

so dividing both sides by $P(B)$ gives the above theorem.

Bayes Example

- N.B. The numbers in this example are made up
- Zika virus effects 1/10,000 people in Puerto Rico, and after returning from vacation, you decide to get tested. The test has a false positive rate of 2%. Your test comes back positive. What is the probability that you have Zika?

Session Objective

1. Define fundamental laws of:
   set notation
   combinatorics
   probability

2. Use these laws to solve problems!