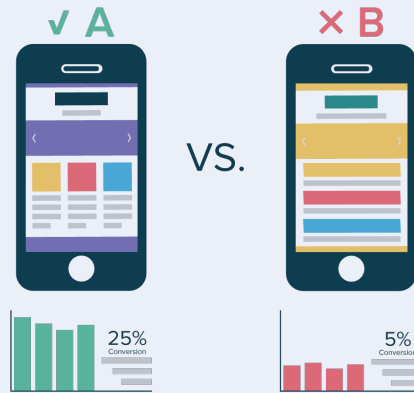
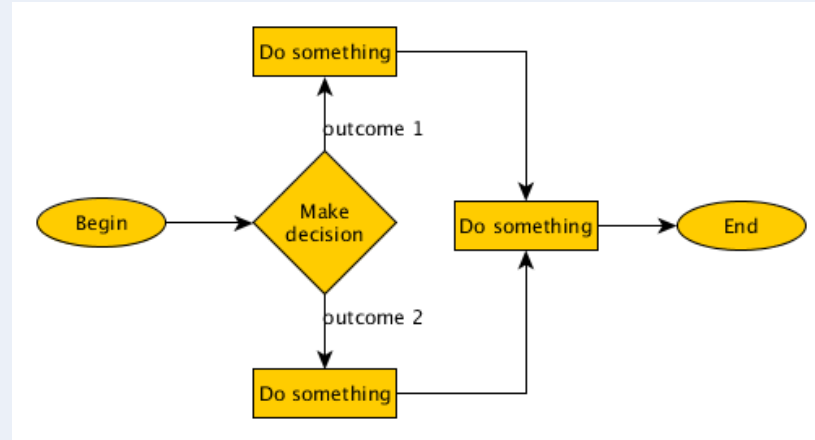


Applications of Statistics in Data Science

A/B Testing

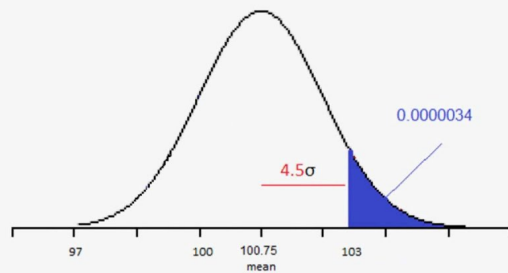


Machine Learning Engineer



Manufacturing

Even if the mean of the production, in the long run, shifts by 1.5σ to 100.75mm, the chance of products being produced out of specification limitations, longer than 103mm, will be less than 3.4 in million.



Clinical Trials

Why do we need statistics for clinical trial design?

- A clinician said: "I try a new drug on my patients. If I like it, I like it."
- The problem is that sometimes you don't know whether you like it or not.

Table: # of successes / total # of patients

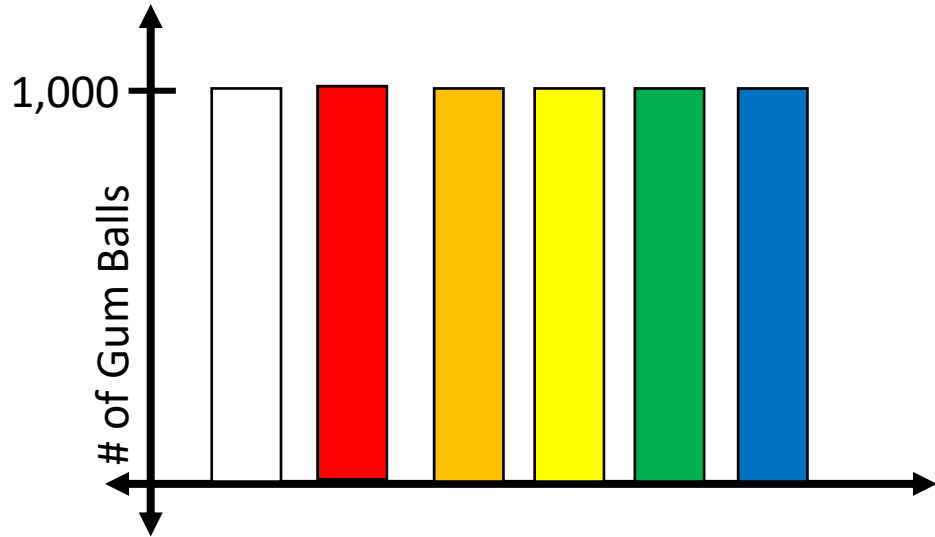
Treatment A	Treatment B
2/5	3/5
4/10	6/10
20/50	30/50
40/100	60/100

Question: when is good enough for you to like it?

Probability vs. Statistics

Probability

Count every gum ball

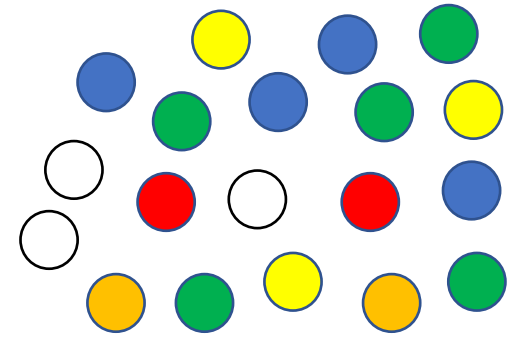


Use the exact distribution to compute properties

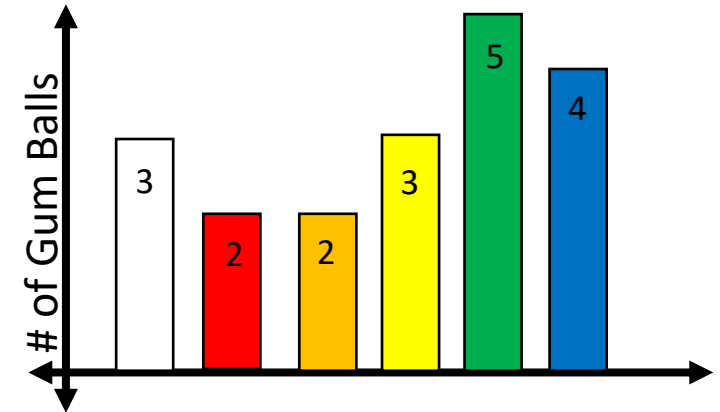


Statistics

Draw Sample



Make Inference



Fisher's Tea Experiment



Can Muriel tell the difference between milk poured into tea and tea poured into milk?

milk poured into tea



tea poured into milk



Probability vs. Statistics

Bayesian

Belief:

I believe that there is a 90% chance that the coin is fair

+

Data:



=

$P(\text{fair} | \text{data}) = 95\%$



Frequentist

Get Data:



Make Inference

Null: My coin is fair

Alternative: My coin is not fair

Conclusion: I fail to reject my null hypothesis