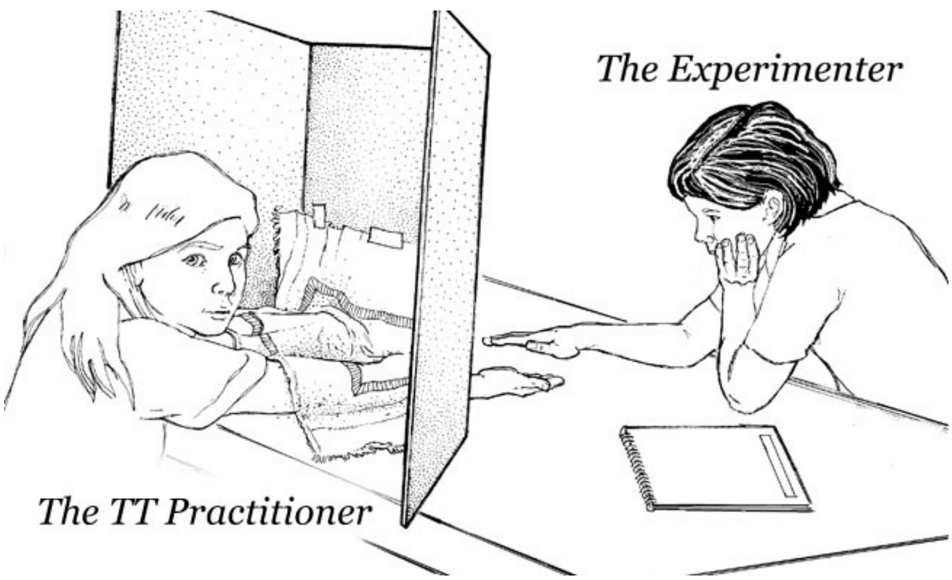
Lecture 1 - Review of classical inference

Guiding example: testing claims of therapeutic touch

↳ google "Emily Rosa"

↳ practitioners claim to feel "Human Energy Field"

↳ Rosa's experiment: experimenter randomly selects one of the practitioner's hands and holds her hand above it. (all occluded behind screen) Practitioner is asked to identify which hand the experimenter has selected.

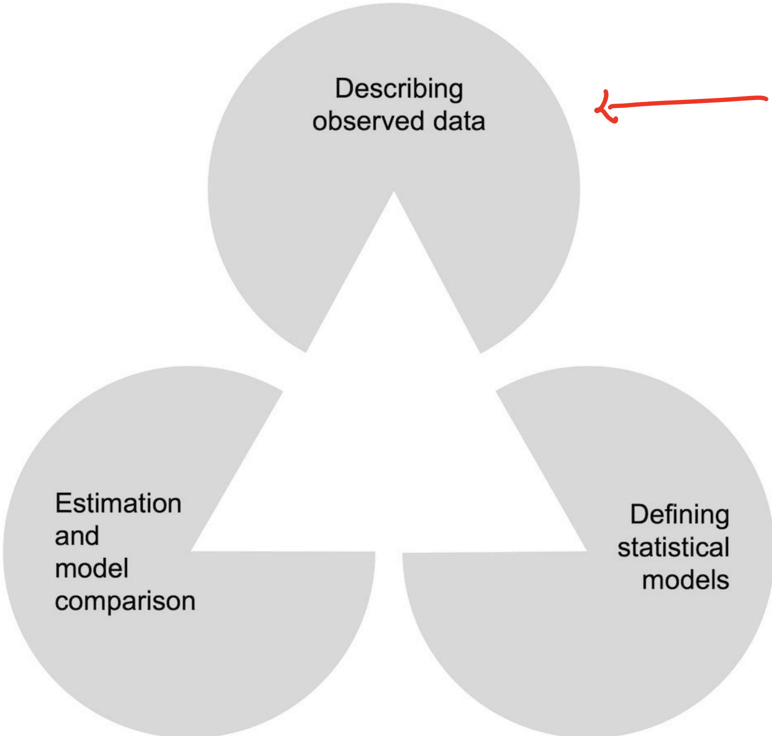


↳ Results: practitioners only correct on 44% of trials.

↳ Goal: evaluate practitioners' claims

Framework of inferential statistics (from Faulkenberry, 2022)

correct choice on 44% of trials



Two methods:

(1) classical

(2) Bayesian

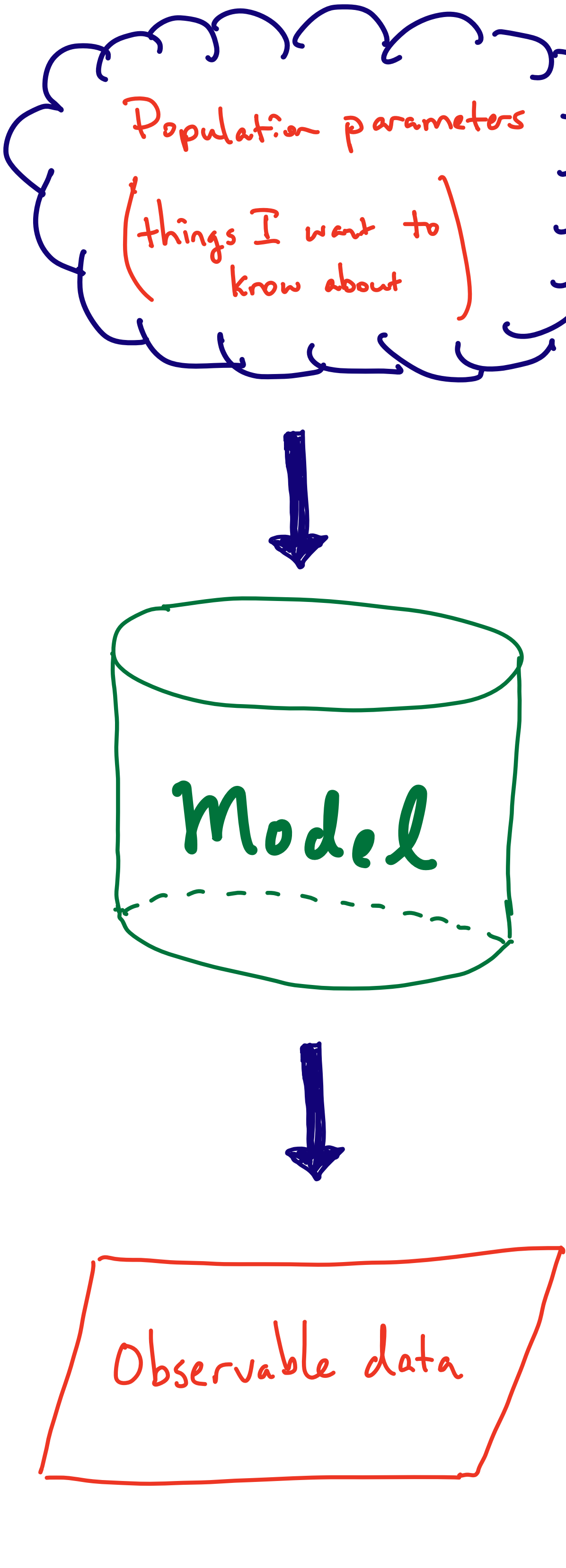
will review in this lecture

What is a model?

↳ mathematical function that links some latent (population) variables to observable data.

probability distribution

Visualization:



Probability of getting a success on a trial (w)

Binomial model

Number of successes out of a set of trials (x)

Binomial model - for a given probability of success ω, the binomial model gives probability of observing x successes across N trials.

Example: Suppose $N = 10$ trials, probability of success on each trial is $\omega = 0.25$. What is probability of observing $x = 8$ successes?

* 10 trials must be either Success or Failure.

↳ 8 Successes → 2 Failures

Observe: S S S S S S S S F F

Prob: $\left[0.25 \cdot 0.25 \cdot 0.25 \cdot 0.25 \cdot 0.25 \cdot 0.25 \cdot 0.25 \cdot 0.25\right] \cdot \left[0.75 \cdot 0.75\right]$

$$= (0.25)^8 \cdot (0.75)^2$$

But, there are multiple ways to get 8 Successes & 2 Failures.

Ans: $\binom{10}{8}$ = "10 choose 8" = $\dfrac{10!}{2! \cdot 8!}$

$$= \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \quad 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 45$$

So $P(x=8) = \binom{10}{8} \cdot (0.25)^8 (0.75)^2$

$= 0.000386$

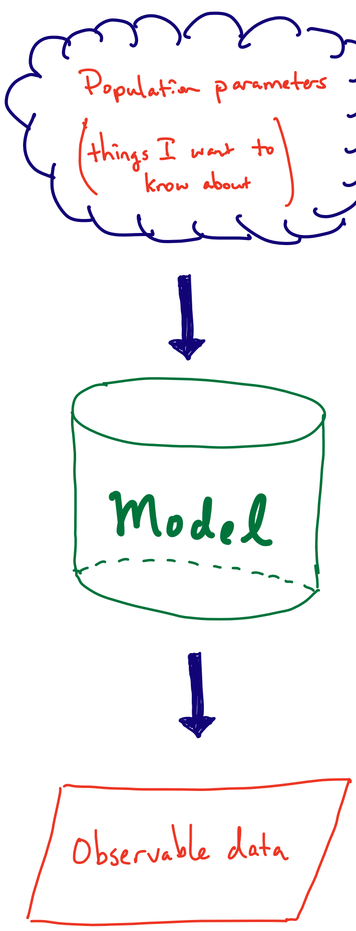
(Not very likely!)

Can do these computations in R!

```r
# Example: w = 0.25, N = 10 trials, probability of 8 successes?
choose(10,8) * (0.25)^8 * (0.75)^2

# R has a built-in function for this!
dbinom(x=8, size=10, prob=0.25)
```

How does the binomial model work as a model?

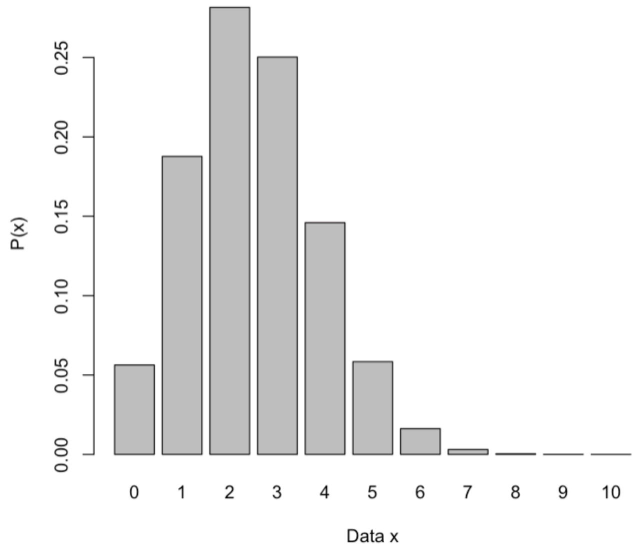


Start with a value for w (prob of success on each trial)

Generate potential data that could be observed.

Plotting the potential data

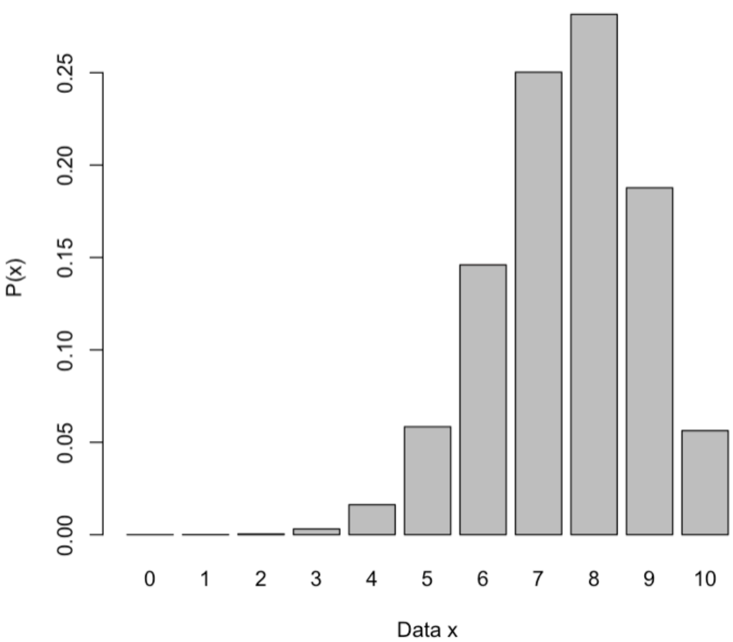
```r
x = seq(from=0, to=10, by=1)
barplot(dbinom(x, size=10, prob=0.25),
        names.arg = 0:10,
        xlab="Data x",
        ylab="P(x)")
```



* small numbers of successes more likely.

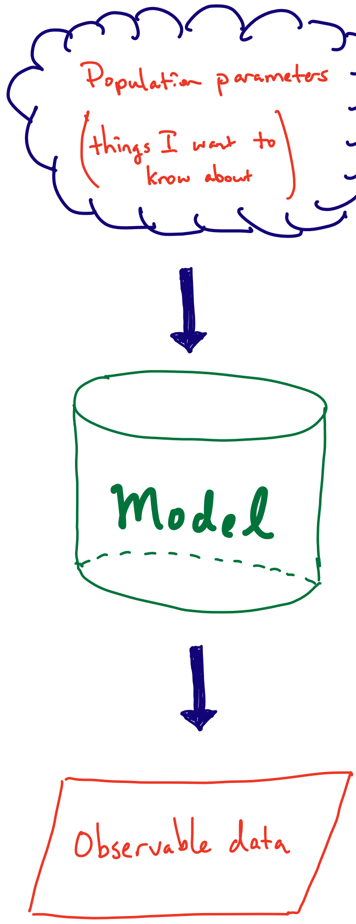
```r
x = seq(from=0, to=10, by=1)
barplot(dbinom(x, size=10, prob=0.75),
        names.arg = 0:10,
        xlab="Data x",
        ylab="P(x)")
```

← only change the parameter ω.



* large numbers of successes more likely.

Problem of inference – given data, what is parameter?



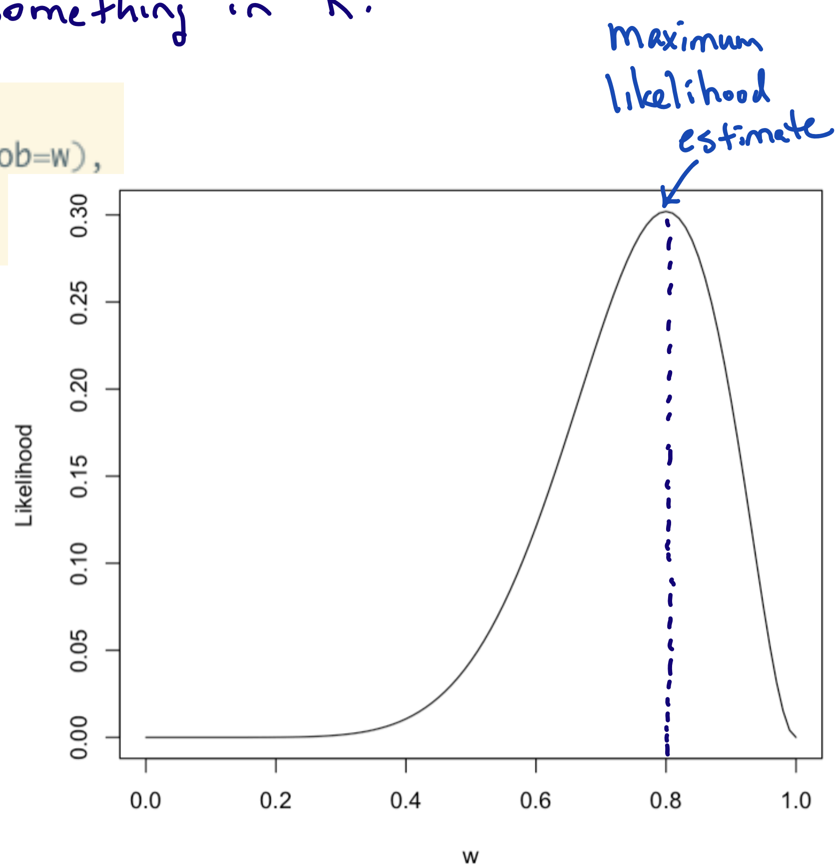
What is the value of the parameter w?

Suppose we observe $x=8$ successes in 10 trials.

To solve this, let's plot something in R:

```r
w = seq(from=0, to=1, by=0.01)
plot(w, dbinom(x=8, size=10, prob=w),
     type="l",
     ylab="Likelihood")
```

"Likelihood function"



Problem of inference

↳ given data, what is/are the value(s) of the model parameter(s)?

↳ from above, we can see that there is uncertainty with respect to these parameter values

↳ Bayesian statistics provides a comprehensive way to quantify this uncertainty.

↳ more in Lecture 2.

As a review, let's talk about model comparison (hypothesis testing)

Step 1: set up two models / hypotheses

* Null hypothesis - performance based on "guessing"

↳ $\mathcal{H}_0 : w = 0.5$

* Alternative hypothesis - performance better than guessing

↳ $\mathcal{H}_1 : w > 0.5$

Step 2 - compare the models

↳ which one best predicts the observed data?

↳ classical method (due to R. Fisher, ~1920):

* Assume $\mathcal{H}_0$ is true. Then calculate the probability of observing the data (or more extreme) under this assumption.

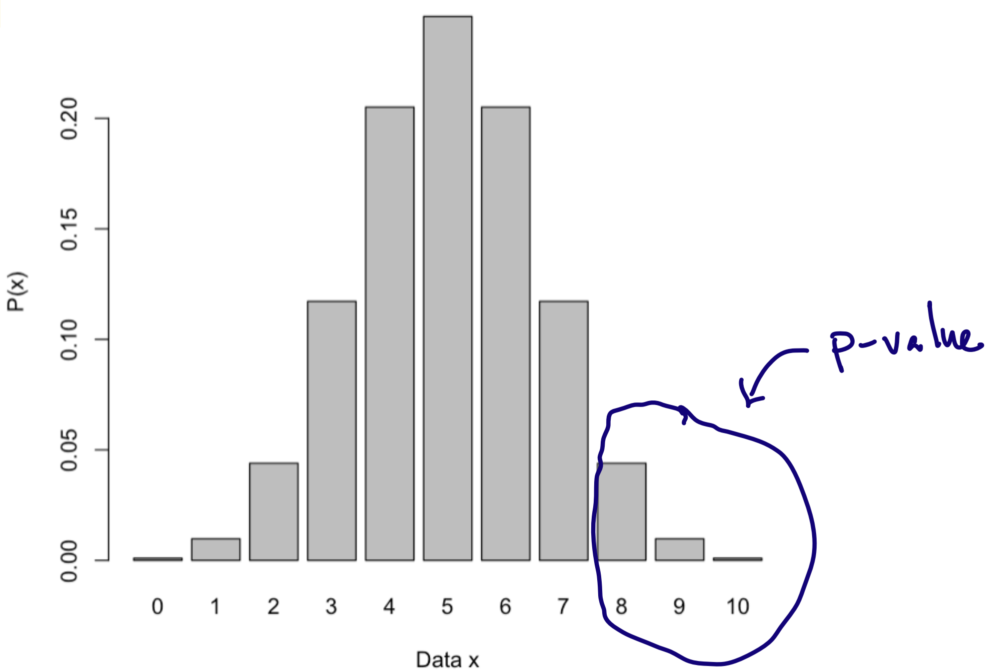
* compute $P(\text{data} \mid \mathcal{H}_0)$

$$= P(x \geq 7 \mid w = 0.5)$$

$$= \text{"p-value"}$$

Let's use R:

```
x = seq(from=0, to=10, by=1)
barplot(dbinom(x, size=10, prob=0.5), # change to 0.5 to reflect null
        names.arg = 0:10,
        xlab="Data x",
        ylab="P(x)")
```



Computing this p-value is easy in R

↳ "pbinom" function – computes prob ≤ given data.

↳ since we want prob ≥ data, must invert it and adjust (1 – pbinom)

```
53  1 - pbinom(7, size=10, prob=0.5)
```

= 0.0547

This p-value is small (i.e, $p = 0.05$) which means that the observed data is rare under H_0.

Thus, we reject H_0 as a plausible model, leaving us with support for H_1.

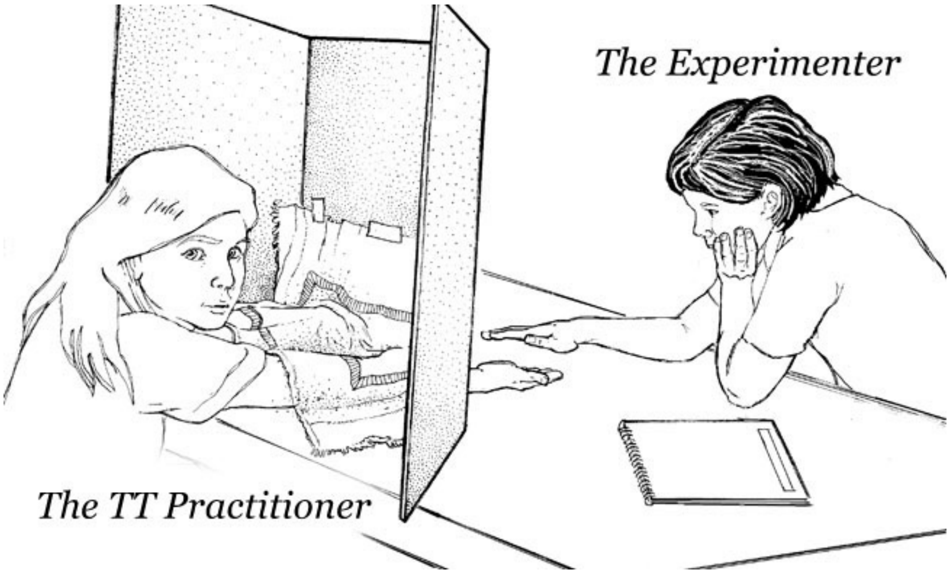
Note: that was a conceptual way to perform the hypothesis test. R will do it directly!

```
> binom.test(x=8, n=10, p=0.5, alternative="greater")

	Exact binomial test

data:  8 and 10
number of successes = 8, number of trials = 10, p-value = 0.05469
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.4930987 1.0000000
sample estimates:
probability of success
                   0.8
```

Back to the Emily Rosa study:



↳ Results: practitioners only correct on 44% of trials.
(actual data from JAMA paper – 123 out of 280 trials).

$H_0: w = 0.5$ (practitioners were guessing)

$H_1: w > 0.5$ (practitioners performing above chance)

```
> binom.test(x=123, n=280, p=0.5, alternative="greater")

	Exact binomial test

data:  123 and 280
number of successes = 123, number of trials = 280, p-value = 0.9819
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.3893703 1.0000000
sample estimates:
probability of success
             0.4392857
```

observed data very plausible under H_0.