- The theory behind (some of) the stuff we covered in Prob Stats,
- Estimators (statistics) — best guess for a population parameter
- In probability we assumed we had μ, σ. But what if we don't?
- And what are the distributions of these boyos?
- What's the distribution of $\bar{x}$? $\bar{x}^2$?
- So use the estimators, distributions to do hypothesis testing and create conf. intervals
- 1st half of this course is the hard part, then it gets easier for some reason
- Squiggle boi
- What is statistical modeling? Why do we care?
- The scenario: Measure the distance from your home to Bunker
- How? Ruler, car's odometer, Google Maps
- How many times? Measure just once? Several times?
- Let $\mu =:$ true distance
- n measurements, $x_1, \ldots, x_n$
- Does $x_i = \mu$ all the time? Nope. $x_i = \mu + \epsilon_i$, where ϵ_i is some error term
- Probably not always positive or always negative, generally some of both, unless the system is biased
- ASSUMPTIONSZ about ϵ_i:
    1. Errors have mean 0
    2. Errors are independent
    3. Errors are symmetric about the mean
    4. All errors have the same variance
- Let's make a model! Assume $\epsilon_1, \ldots, \epsilon_n$ are independent, distributed $N(0, 1)$
- Question: What is the distribution of the x_is? They're normal! That's because you are just adding a const μ to $N(0, 1)$
- So what's $E(x_i), V(x_i)$?
- $E(x_i) = E(\mu + \epsilon_i) = \mu + E(\epsilon_i) = \mu + 0 = \mu$
- $V(x_i) = V(\mu + \epsilon_i) = V(\epsilon_i) = 1$
- $\Rightarrow x_1, \ldots, x_n$ are IID $N(\mu, 1)$

- We have two different sets of RVs. The xs, $x_1, \ldots, x_n$ are the observables, the measurements, the guys, the those dudes
- Also RVs in terms of the ϵs, $\epsilon_1, \ldots, \epsilon_n$, and they're unobservable
- μ is an unknown parameter, let's estimate it
- What we need to know: "best" estimator for μ? $\bar{x}$, maybe the mean?
- Estimator = best guess for the parameter
- An estimator is just a statistic
- Recaaaaaaaall: What's the CDF? It's the probability that an RV $X \leq x$ for some value of $x, F(x) = P(X \leq x)$
- For jointly distributed RVs: $F(X_1, \ldots, X_n) = P(X_1 \leq x_1 \cap \ldots \cap X_n \leq x_n)$
- Big ol cappy boi. It's an intersection
- The latter part of 334 is looking at dependent RVs and stuff
- But that's not what we'll be doing. We're dealing with independent boyos only! It's not too unreasonable to assume samples are IID
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This is always true
- If A, B are independent, then $P(A \cap B) = P(A)P(B)$
- Joint CDF is much easier to compute if they're IID
- If independent, joint CDF is $\prod_{i=1}^{n} F(x_i)$
- Paaaaarmetric Moooodeeeeeeeeeelz:
- Let M be a model for $x_1, \ldots, x_n$. If all CDFs of M can be gotten by varying one or more params, then M is a parametric model
- Let M be a parametric model, param θ. Set of all possible values for $\theta =$: the parameter space
- Where is your parameter allowed to exist such that a PDF or PMF exists? Anywhere such that it integrates to 1, and there is no point that the function is less than 0 on $(-\infty, \infty)$
- Ex: $x \sim Bin(n, p), n$ is known, p is probability of success $0 \leq p \leq 1$
- $X \sim Poisson(\lambda)$, exists if $\lambda > 0$
- Statistic: Something that describes the data. A single point. A function of observable RV (data) where it does not depend on any unknown params
  1. $\bar{X}$: Yep, a statistic

2. $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$ Yep
3. σ^2 Nope, you can't get it from only a sample, we don't know σ
4. $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ Nope, we don't have μ or σ
5. Min and max of the X guys? Yep, a statistic