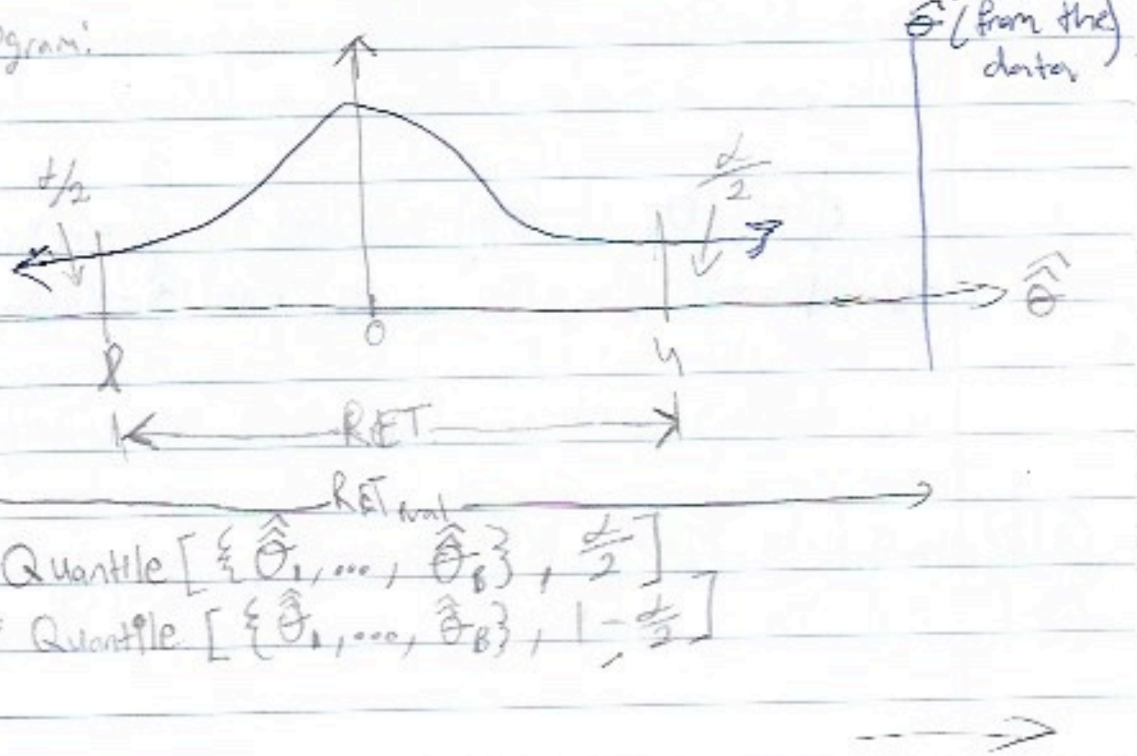
Lec 22

Let's go with (a) the difference in sample averages. So, under the null, $\bar{x}_{D,1} - \bar{x}_{D,2} \approx 0$. It won't be exactly zero due to chance variations, so what is the threshold of non zero to reject H_0? We need the sampling distribution of this difference in sample averages under H_0.

There are many such fake samples $\binom{200}{100} \approx 10^{58}$. So w/ a big computer, let's take $B = 1{,}000{,}000$. The higher the B is the more accurate, but not much more accurate. Let's calculate the test statistic for all B "resamplings" and plot a histogram:



$\ell := \text{Quantile}\left[\{\hat{\hat{\theta}}_1, \ldots, \hat{\hat{\theta}}_B\}, \frac{\alpha}{2}\right]$

$u := \text{Quantile}\left[\{\hat{\hat{\theta}}_1, \ldots, \hat{\hat{\theta}}_B\}, 1-\frac{\alpha}{2}\right]$

$\dashrightarrow$

How to get the 1-sided pval. You count the number of $\hat{\theta}_b$'s that are more extreme than the $\hat{\theta}$ from the original data.

You can also uses Fisher's permutation idea to get a CI for θ. This is difficult computationally so we won't study it.

We will now see a resampling method that is extremely famous and extremely useful: Efron's non-parametric bootstrap (1979). ch 8 A.5 and the idea is remarkly simple. Imagine you have an iid DGP $f(x; \theta_1, \ldots, \theta_k)$ and you have some parameter/function of parameters you want inference for e.g.

$$\phi = g(\theta_1, \ldots, \theta_k) \text{ estimated by } \hat{\phi} = w(x_1, \ldots, x_n)$$

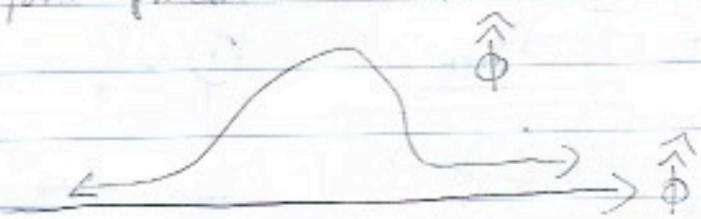
$$\text{e.g. } \phi = \text{med}[X] = g(\theta_1, \ldots, \theta_k) \text{ but } \hat{\phi} = \text{Median}[X_1, \ldots, X]$$

We don't know to get the estimator's distribution. Which means we can't test hypothesis or construct CI's. But... the bootstrap gives us a way to find the asymptotic distribution of the estimator:

For $b = 1, \ldots, B$, a large # of resamplings,
* take a fake sample of size n by sampling w/ replacement of the original n data points (this will give you about 2/3 of the unique data points and 1/3 left out)
* for each fake sample, I calculate $\hat{\theta}_b$.

The collection of these B statistics is called the "bootstrap distribution"

Thm: the bootstrap distribution approximates the real distribution of the estimator as $B \to \infty$ and $n \to \infty$. So B should be as big as possible given your practical constraints.



One way to get a CI is to just use the quantiles of the bootstrap distribution:

$$CI_{\phi, 1-\alpha} = \left[\text{Quantile}\left[\{\hat{\hat{\phi}}_1, \ldots, \hat{\hat{\phi}}_B\}, \frac{\alpha}{2}\right], \text{Quantile}\left[\{\hat{\hat{\phi}}_1, \ldots, \hat{\hat{\phi}}_B\}, 1-\frac{\alpha}{2}\right] \right]$$

there are other ways too that we won't study. One way to do a two-sided hypothesis test (i.e. against $H_0 : \phi = \phi_0$) is to just check

$$\phi_0 \in CI_{\phi, 1-\alpha} \overset{Yes}{\Longrightarrow} \text{Retain } H_0$$

In finance, people care about the Sharpe Ratio defined as:

a constant representing the risk free rate

$$\phi = \frac{E[X] - r_{free}}{SD[X]}$$

(estimate)

$$\hat{\phi} = \frac{\bar{X} - r_{free}}{\sqrt{\frac{1}{n} \sum (X_i - \bar{X})^2}} \qquad \hat{\phi} \sim ?$$

Randomization and Causal Estimation

Consider two populations.

Scenario 1:

Pop. 1: Students who hand in their test at the end of the exam

Pop. 2: Students who hand in their test early.

Let $Y_{1,1}, \ldots, Y_{1,n-1}$ denote test score data from Pop. 1

Let $Y_{2,1}, \ldots, Y_{2,n-2}$ denote test score data from Pop. 2

We want to test if the means of these two populations are diffrent i.e. against $H_0: \theta_1 = \theta_2$. We know nothing about the DGPs but we can use the Wald test (2-sample z-test):

If $\dfrac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \in [-1.96, 1.96] \Rightarrow$ Retain H_0 at $\alpha = 5\%$

Let's say you reject and then you can conclude: "there is a diffrence in the mean test scores b/w students who hand