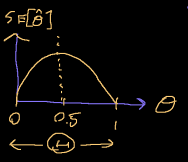
DGP: $X_1, \ldots, X_n \overset{iid}{\sim}$ with mean θ, variance σ^2

If $\hat{\theta} = \bar{X} \Rightarrow \hat{\theta}$ is unbiased $\quad R(\hat{\theta}, \theta) = \frac{\theta(1-\theta)}{n} = MSE$

$$SE[\hat{\theta}] = \sqrt{Var\left[\frac{1}{n}(X_1 + \ldots + X_n)\right]} = \sqrt{\frac{1}{n^2} \sum Var[X_i]} = \sqrt{\frac{1}{n^{\cancel{2}}} \cancel{n} \sigma^2} = \frac{\sigma}{\sqrt{n}}$$

$$= \sqrt{\frac{\theta(1-\theta)}{n}}$$

$X_1, \ldots, X_n \overset{iid}{\sim} Bern(\theta)$



$\sup_{\theta \in \Theta} R(\hat{\theta}, \theta) = \frac{1}{4n}$

supremum (maximum)

Goal #3 of inference: theory testing (hypothesis testing).

You have some well-specified mathematical theory about the DGP. For example, in the iphone survey, "I think the proportion of iPhone users in the population is NOT 52.4%. I want to prove my theory to the world (using my sample).

Note: it is absolutely impossible to prove or disprove my theory because you cannot see the whole population (or go inside of the DGP). We must use inference which is always a guess.

Two ways to go about "proving" my theory:
(1) I assume I'm right and wait for other people to show me data that contradicts my theory.
(2) I assume my theory is wrong. Then I adduce (bring) evidence (i.e. data) to the contrary until people are convinced my theory is right.

#2 is more intellectually honest and more likely to convince.

A "hypothesis" is a mathematical statement about the DGP e.g. theta = 0.9, theta > 0.9, theta is not equal to 0.9, or theta <= 0.9 or theta is in the set [0.89, 0.91], etc.

The "alternative hypothesis" (H_a) is the theory you want to prove. The "null hypothesis" (H_0) is the opposite you assum in #2 for the purpose of contradicting it. Usual cases:

$H_0: \theta \leq \theta_0, \quad H_a: \theta > \theta_0$ (right-tailed test)

$H_0: \theta \geq \theta_0, \quad H_a: \theta < \theta_0$ (left-tailed test)

$H_0: \theta = \theta_0, \quad H_a: \theta \neq \theta_0$ (two-tailed test)

How to perform this test? There are many, many options even for the same DGP. The protocol goes as follows
(1) you think of a "test statistic" that could measure the departure away from H_0.
(2) Derive the statistical estimator's distribution under H_0.
(3) Gauge the departure.

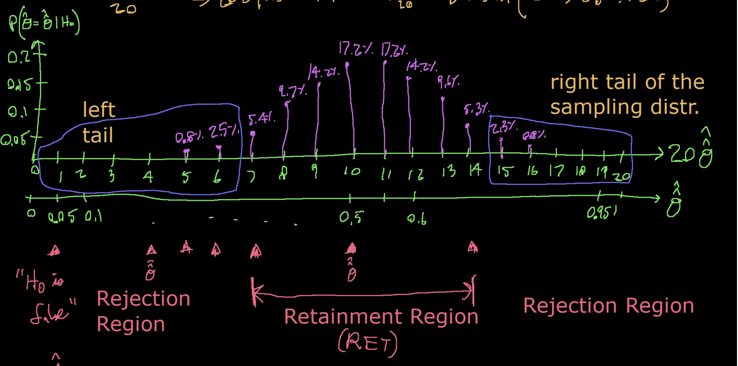
We begin with DGP: iid Bern(theta) and the "binomial exact test"

$H_a: \theta \neq .524^{=\theta_0}, \quad H_0: \theta = .524^{=\theta_0}$

(1) My test statistic is .... $\hat{\hat{\theta}} = \bar{x}$. $\hat{\hat{\theta}}$ is a realization from $\hat{\theta}$.

(2) $\hat{\theta} \,|\, H_0 \sim ?$ $\quad n = 20$

$$\hat{\theta} = \frac{X_1 + \ldots + X_{20}}{20} \Rightarrow 20\hat{\theta} \,|\, H_0 = X_1 + \ldots + X_{20} \sim Binom(20, \theta_0 = .524)$$



$\hat{\hat{\theta}} \in RET \Rightarrow$ Retain H_0 (fail to reject H_0). Not enough evidence to reject H_0. Some authors say "accept H_0".

$\hat{\hat{\theta}} \notin RET \Rightarrow$ Reject H_0 / Accept H_a. My estimate is "statistically significant".

Let's say we rejected H_0 but it was really was true. This is called a Type I error. Where is the P(Type I error) on our plot?

$$\alpha := P(\text{Type I error}) = P(\hat{\theta} \notin RET \,|\, H_0)$$

Then in a 2-tailed test, I apportion about alpha/2 to the left tail and about alpha/2 to the right tail.

In my RET,

$$\alpha = P(\hat{\theta} = 0 \,|\, H_0) + \ldots + P(\hat{\theta} = 0.3 \,|\, H_0) + P(\hat{\theta} = 0.75 \,|\, H_0) + \ldots + P(\hat{\theta} = 1 \,|\, H_0) = 7.06\%$$

The choice of alpha is up to you. The scientific community's standard is 5% and sometimes 1%.

If you would like to prove your theory, you have to accept a positive probability of a Type I error.

If I fail to reject H_0 when H_a is true that's a different error, a "type II error". Failure to prove your theory.

The smaller the alpha, the larger the P(Type II error).

Truth \ Decision	Retain H_0	Reject H_0
H_0	✓	Type I error
H_a	Type II error	✓

As of now, we cannot calculate the P(Type II error).