decision errors

- type 1 & type 2 errors
- balancing error rates

Dr. Mine Çetinkaya-Rundel
Duke University

		Decision	
		fail to reject H_0	reject H_0
Truth	H_0 true	✔	Type 1 error
	H_A true	Type 2 error	✔

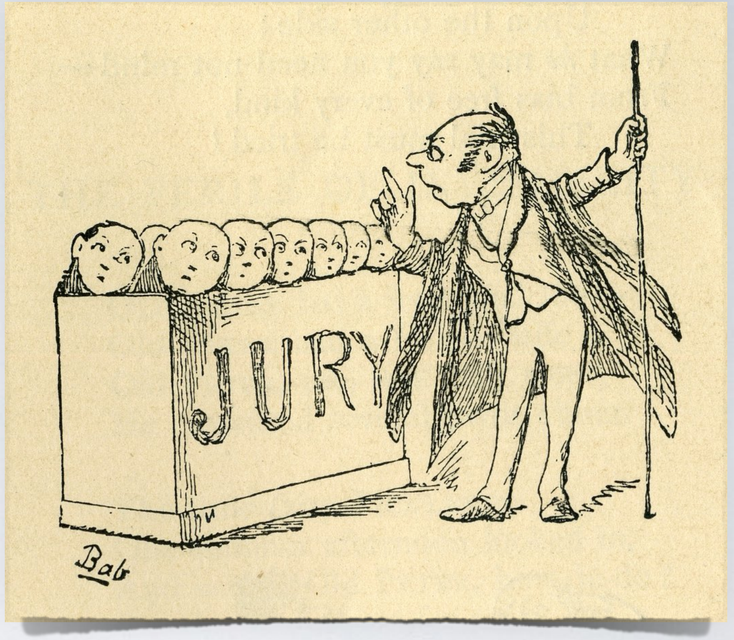
- Type 1 error is rejecting H_0 when H_0 is true.
- Type 2 error is failing to reject H_0 when H_A is true.
- We (almost) never know if H_0 or H_A is true, but we need to consider all possibilities.

hypothesis test as a trial

If we again think of a hypothesis test as a criminal trial then it makes sense to frame the verdict in terms of the null and alternative hypotheses:

H_0 : Defendant is innocent

H_A : Defendant is guilty

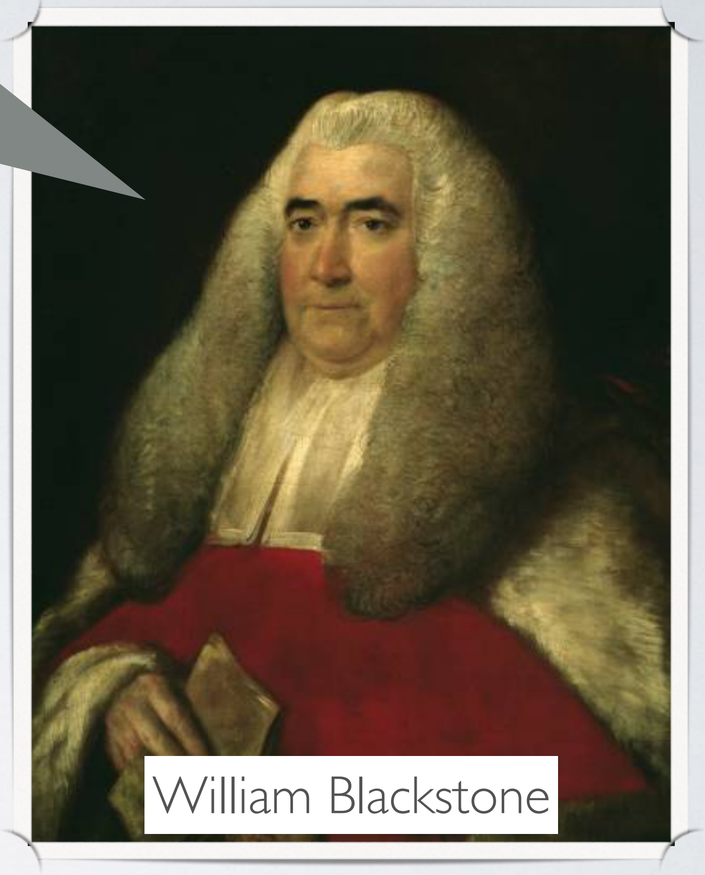


Which type of error is being committed in the following circumstances?

- Declaring the defendant innocent when they are actually guilty *Type 2 error*
- Declaring the defendant guilty when they are actually innocent *Type 1 error*

Jury: http://upload.wikimedia.org/wikipedia/commons/5/5d/Trial_by_Jury_Usher.jpg

"better that ten guilty persons escape than that one innocent suffer"


William Blackstone

Which error is the worst error to make?

- Type 2 : Declaring the defendant innocent when they are actually guilty
- Type 1 : Declaring the defendant guilty when they are actually innocent

William Blackstone: http://en.wikipedia.org/wiki/File:SirWilliamBlackstone.jpg.

type I error rate

- We reject H_0 when the p-value is less than 0.05 ($\alpha = 0.05$).
- This means that, for those cases where H_0 is actually true, we do not want to incorrectly reject it more than 5% of those times.
- In other words, when using a 5% significance level there is about 5% chance of making a Type I error if the null hypothesis is true.

$$P(\text{Type I error} \mid H_0 \text{ true}) = \alpha$$

- This is why we prefer small values of α – increasing α increases the Type I error rate.

choosing α

If Type 1 Error is dangerous or especially costly, choose a small significance level (e.g. 0.01).

Goal: we want to be very cautious about rejecting H_0, so we demand very strong evidence favoring H_A before we would do so.

If a Type 2 Error is relatively more dangerous or much more costly, choose a higher significance level (e.g. 0.10).

Goal: we want to be cautious about failing to reject H_0 when the null is actually false.

Scale: http://commons.wikimedia.org/wiki/File:US_Department_of_Justice_Scales_Of_Justice.svg

goal:
keep α and β
low

		Decision	
		fail to reject H_0	reject H_0
Truth	H_0 true	$1 - \alpha$	Type 1 error, α
	H_A true	Type 2 error, β	$1 - \beta$

- Type 1 error is rejecting H_0 when you shouldn't have, and the probability of doing so is α (significance level).
- Type 2 error is failing to reject H_0 when you should have, and the probability of doing so is β.
- Power of a test is the probability of correctly rejecting H_0, and the probability of doing so is $1 - \beta$

type 2 error rate

If the alternative hypothesis is actually true, what is the chance that we make a Type 2 Error, i.e. we fail to reject the null hypothesis even when we should reject it?

- The answer is not obvious.
- If the true population average is very close to the null value, it will be difficult to detect a difference (and reject H_0).
- If the true population average is very different from the null value, it will be easier to detect a difference.
- Clearly, β depends on the effect size (δ), difference between point estimate and null value.