Stat 201:
Introduction to Statistics

Standard 29: Types of Errors

Errors Associated with Testing

- **Type I Error** occurs when H_o is rejected but in reality H_o is true
  - P(Type I Error) = significance level = 1-confidence level
- **Type II Error** occurs when H_o is not rejected but in reality H_o is false

	H_o *is true*	H_a *is true*
We failed to reject	No Error	Type II Error
We rejected	Type I Error	No Error

Errors Associated with Testing

Implication:

- **Type I Error** occurs when H_o is rejected but in reality H_o is true
    - P(Type I Error) = significance level = 1-confidence level
- If we were to do 1000 hypothesis tests at 95% confidence or, at the .05 level of significance, we would make type one error 5% of the time.
    - – THAT'S 50 TIMES!

Errors Associated with Testing

- This is actually my research area – trying to come up with better methods to do many hypothesis tests which is called multiple testing, or multiplicity.

- This is really important in comparing many different things – particularly the gene expression.
  – Yes, I model genes.

Controlling Error

- **We can control the probability of Type I Error** by our choice of the significance or confidence level
- Though we **can’t control the probability of Type II** Error directly, when we decrease the probability of Type I Error the probability of Type II Error increases

Controlling Error

- **To increase Type I error:** decrease confidence
- **To decrease Type I error:** increase confidence

- **To increase Type II error:** decrease Type I error
  → increase confidence
- **To decrease Type II error:** increase Type I error
  → decrease confidence

Example

- Data is collected to see if there is evidence, at the .1 level of significance or 90% confidence, that the average contaminant concentration level exceeds the acceptable level for fishing, let's call this level μ_0.

- $H_0: Contaminant\ levels\ are\ low: \mu \leq \mu_0$
- $H_a: Contaminant\ levels\ are\ too\ high:\ \mu > \mu_0$

Example

- $H_0: Contaminant\ levels\ are\ low: \mu \leq \mu_0$
- $H_a: Contaminant\ levels\ are\ too\ high: \ \mu > \mu_0$

- **Type I Error:** The committee determines that the contaminant level does exceed the acceptable level incorrectly
  - People can't fish when really it's safe

Example

- $H_0: Contaminant\ levels\ are\ low: \mu \leq \mu_0$
- $H_a: Contaminant\ levels\ are\ too\ high:\ \mu > \mu_0$

- **Type II Error:** The committee determines that the contaminant level doesn’t exceed the acceptable level incorrectly
  – People can eat contaminated fish leading to possible sickness or death

Example

- $H_0: Safe\ to\ fish$
- $H_a: Not\ safe\ to\ fish$

- **Type I Error:** We say it's unsafe when really it is
- **Type II Error:** We say it's safe when really it isn't

Example

- Type I Error: People can’t fish when really it’s safe
- Type II Error: People can eat contaminated fish leading to possible sickness or death

- We see here that **Type II error would be worse** as they would be putting humans in danger. Clearly, it would be important here to **make the probability for Type II Error as small as possible**.
- We’re happy that the **confidence is low** because that indicates a **higher P(Type I Error)** and subsequently a **lower P(Type II Error).**

Example

- If it’s such an important problem why don’t we just shoot for the right answer?

- **We are shooting for the right answer, but in statistics we’re never 100% sure – we only have evidence up to a point!**

Example

- The idea with confidence intervals and statistics is that we’re coming up with an interval or statement about something so big we cannot measure it
- It’s usually easy to get sample measurements, but a lot of times there are too many experimental objects in a population to make if feasible to take measurements for every one of them.

Example

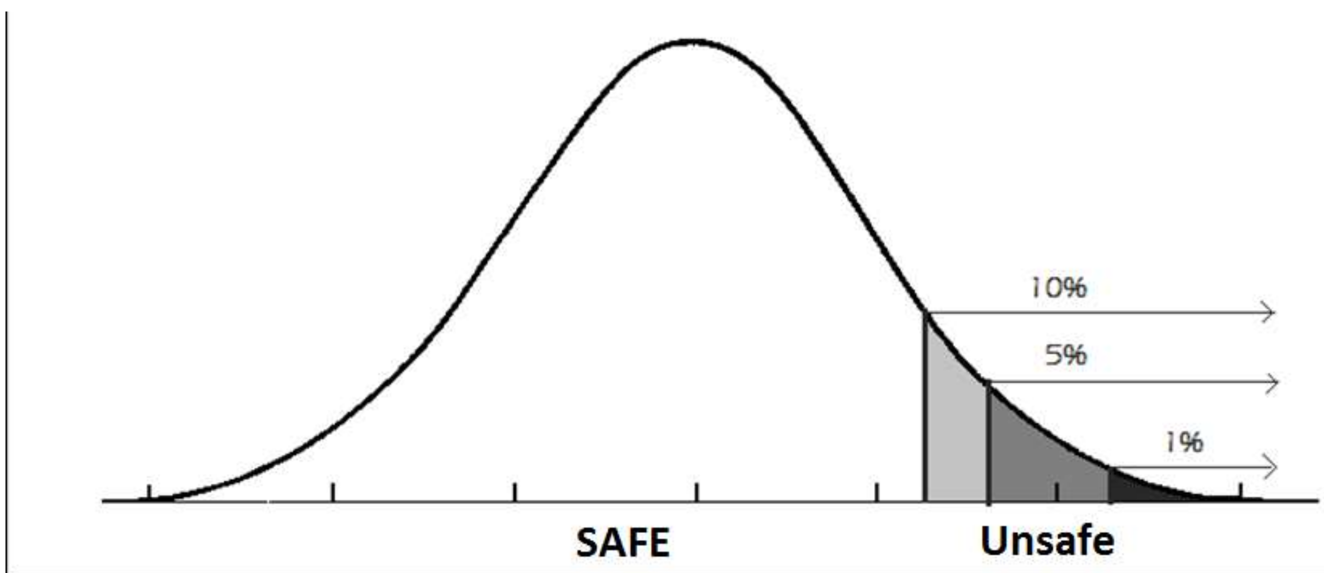
- For instance, according to, , there are over 300 million cubic miles of water on earth!
  – That’s enough to fill over 350 quintillion gallon containers!
- Now, think about taking each one of those gallons of water and testing them for contaminants
  – not bad if it’s just a couple, but you’d be dead before you could check 350 quintillion

Example

- What scientists do is take measurements from random locations around a fishing area and they use statistics to come up with a decision.
- Yes, it would be better to drain all the water and test every molecule but that would be expensive and take an unreasonable amount of time

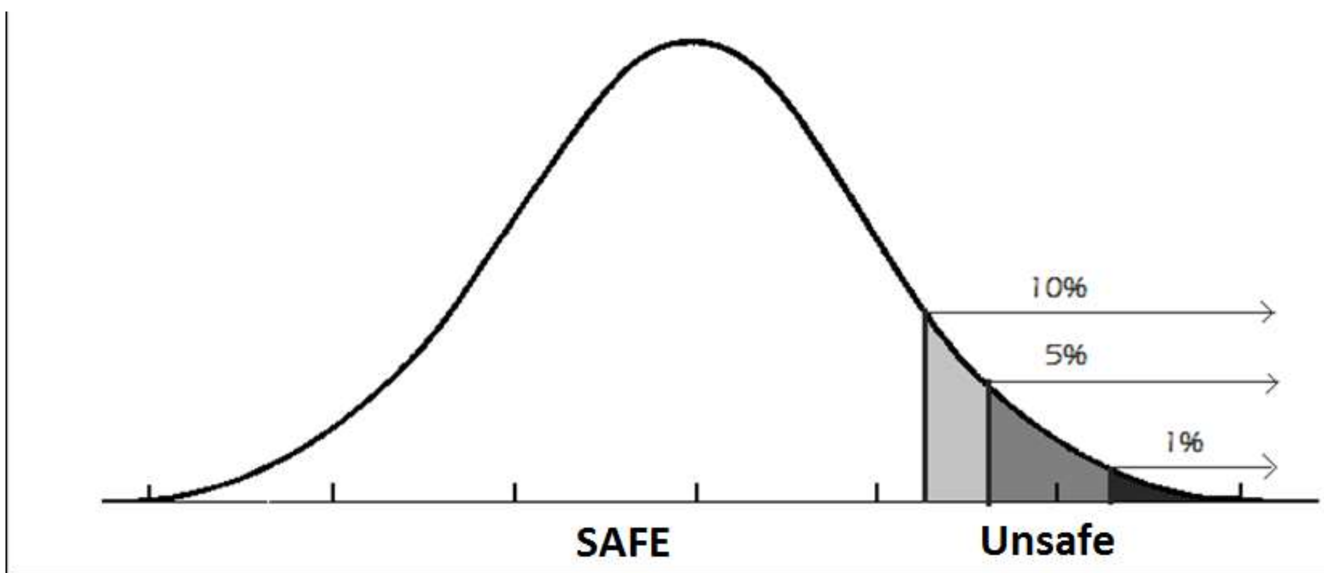
Example

- As we increase our confidence, we make our rejection region smaller which means we’re less likely to reject.
  - We would need a more unusual sample to reject



Example

- As we decrease our confidence, we make our rejection region larger which means we’re more likely to reject.
  - We would need a less unusual sample to reject here



Example

- We're still going for the correct answer, but how we choose our confidence level allows us "sort of choose" what error we'll make if we are incorrect
  - I say sort of choose because it's still possible to make both mistakes – we can only make one mistake more likely than the other