Mech 305-305
Tutorial 3

Hooman Esfandiari
PhD Candidate, Biomedical Engineering

Winter 2019

Hypothesis Testing

Inferential statistics: the process of using data analysis from a sample dataset to infer characteristics of the main population.

Z statistics:

Sample mean

Population mean

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Sample STD

Sample size

Hypothesis Testing

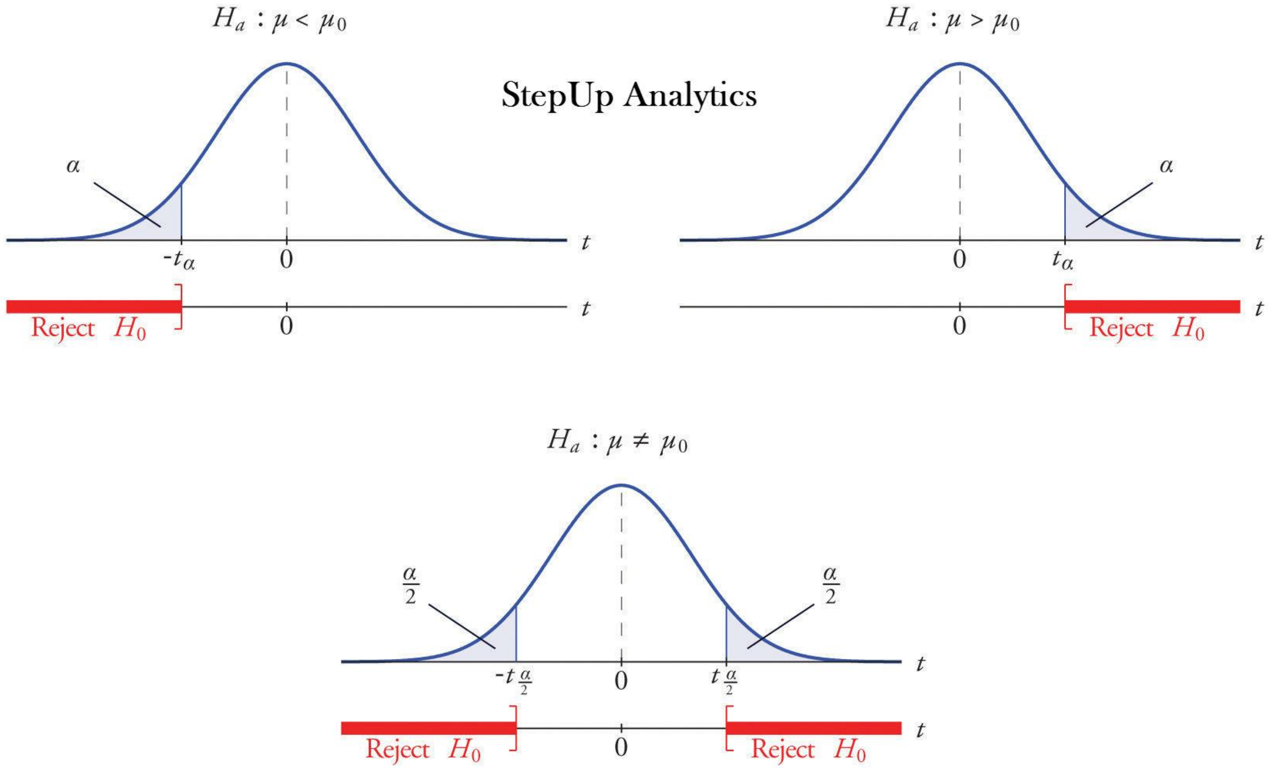


Image courtesy of: stepupanalytics.com

Question 1

$X = [5.5, 5.0, 4.9, 5.6] \rightarrow$ Your measurements

What is the likelihood (alpha value) that the measurements from your test method could suggest a different size even in the case when the bolt diameter actually was 5mm?

1. Form your hypothesis:

$H_0 \rightarrow$ Null hypothesis: $\mu = 5$
$H_1 \rightarrow$ Alternative Hypothesis : $\mu \neq 5$

Question 1(see the board)

2. Calculate the Z values

3. Lookup Z table for critical values

4. Calculate the alpha value (type 1 error)

Bayes Theorem

Imagine you go to a doctor test for a horrible and rare disease that affects about 0.1% of the population in your community → you do the lab work and the test is POSITIVE!

This is how good the test is:

Will correctly identify the 99% of the people who have the disease → True positive

Only incorrectly identify 1% of the people who don't have the disease --> False positive

What are the chances that you actually have this horrible disease?

A) 99% **B) 1%** **C)100%** **D)9%**

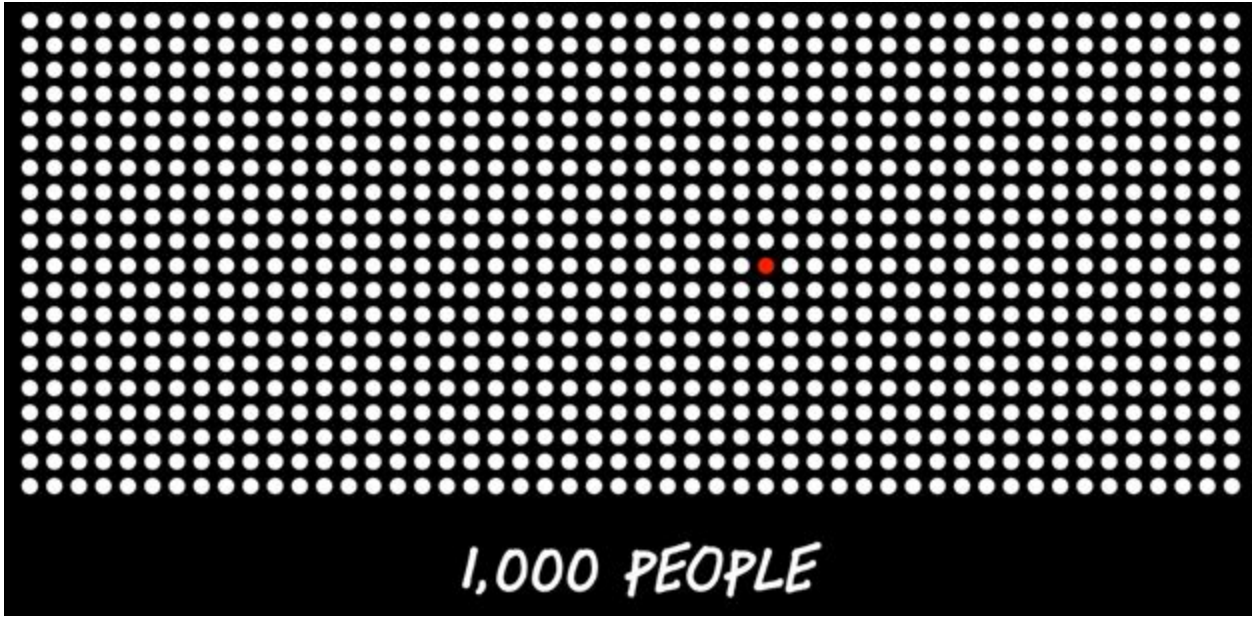
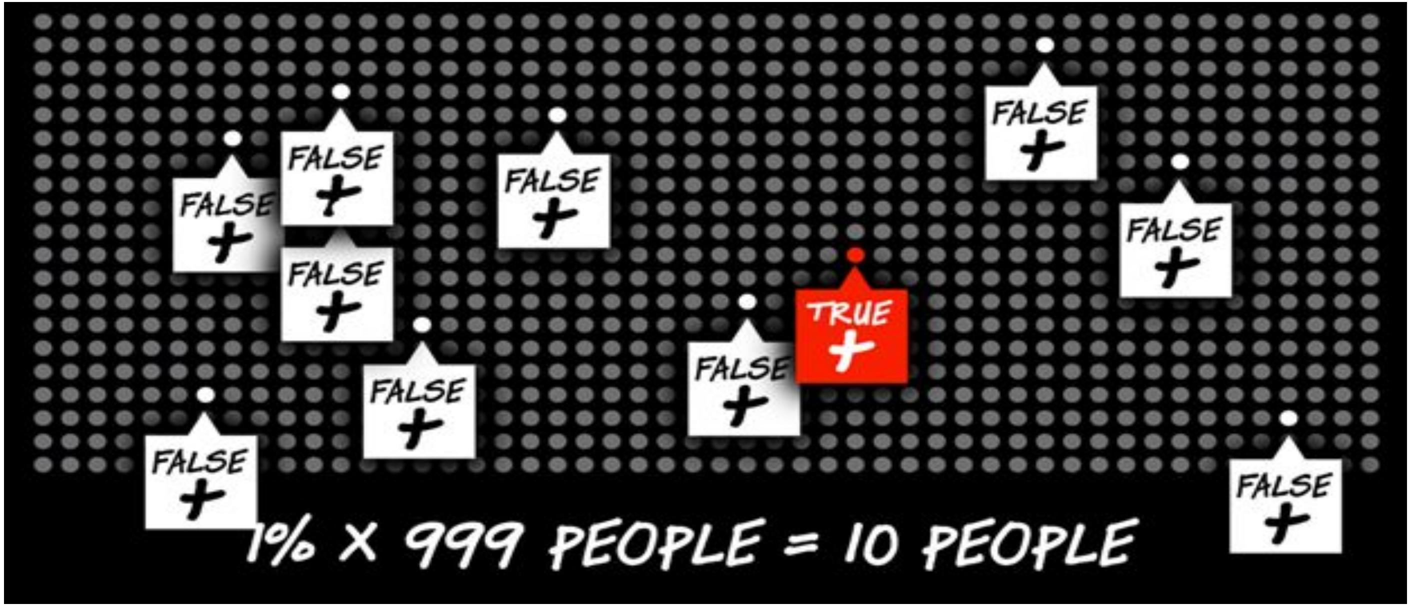


Image courtesy of: Veritasium Youtube Channel


FALSE +
FALSE +
FALSE +
FALSE +
FALSE +
FALSE +
FALSE +
TRUE +
FALSE +
FALSE +
FALSE +
1% X 999 PEOPLE = 10 PEOPLE


1 IN 11 PEOPLE = 9%

$$P(\overset{\text{actually have the disease}}{H} \overset{\text{given}}{\mid} \underset{\text{you tested +}}{E}) = \frac{\overset{\text{you would test + if you had the disease}}{P(E \mid H)} \quad * P(\overset{\text{PRIOR probability of having the disease}}{H})}{\underset{\text{probability of testing +}}{P(E)}}$$

$$P(E) = P(H) * P(E \mid H) + P(-H) * P(E \mid -H)$$

Question 2

Your measurements: X = [5.5,5, 4.9, 5.6]

Actual diameters: ϕ = [3, 4, 5, 6, 7, 8, 10]

Now to simplify the question, let's assume we have made one observation X = [5] and we know there exists two types of bolts: ϕ = [5 (6 of them), 6 (200 of them)]

Remember based on Bayes rule:

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}$$

Question 2

Note that:

$$P(x|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) dx$$

Now plug in the probabilities:

$$P(\theta = 5|x = 5) = \frac{P(x = 5|\theta = 5)P(\theta = 5)}{P(x = 5|\theta = 6) + P(x = 5|\theta = 5)}$$

Question 2

Also note that:

$$P(\theta|x) \sim P(x|\theta)P(\theta)$$

And:

Given that x1 and x2 are two independent variables:

$$P(x_1, x_2) = P(x_1) \times P(x_2)$$

Therefore:

$$P(\theta|x_1, x_2) \sim P(x_1|\theta)P(x_2|\theta)P(\theta)$$

Question 2) Deliverables

Write a matlab script to compute and plot the prior and posterior probabilities for different bolt sizes given the measurements for four cases:

a) priors as given above, σ =0.6 (base case),

b) priors as given above, σ =0.1

c) priors as given above, σ =10

d) No prior information about bolt frequency, σ =0.6

Explain why the measurement uncertainty and prior information affect the calculated most likely bolt size.

Matlab Scripts

1. Create your variables:

    Theta (known characteristics)

    Prior probabilities → probability for each theta

    Measurements (observed quantities)

Matlab Script

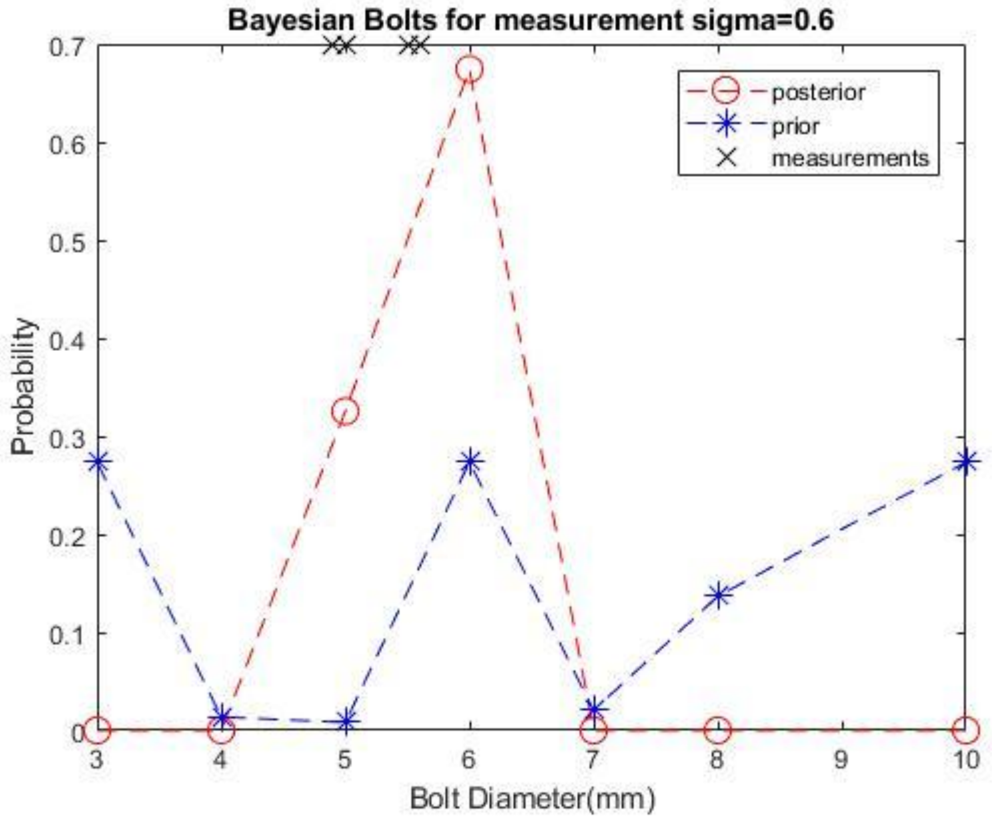
2. Calculate the posterior probabilities individually:

   For example, if Θ=5:

$$P(\theta = 5|x_1, x_2, x_3, x_4) \sim$$

$$P(x_1|\theta = 5)P(x_2|\theta = 5)P(x_3|\theta = 5)P(x_4|\theta = 5)$$

$$\times P(\theta = 5)$$


Bayesian Bolts for measurement sigma=0.6
posterior
prior
measurements
Probability
Bolt Diameter(mm)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
3
4
5
6
7
8
9
10