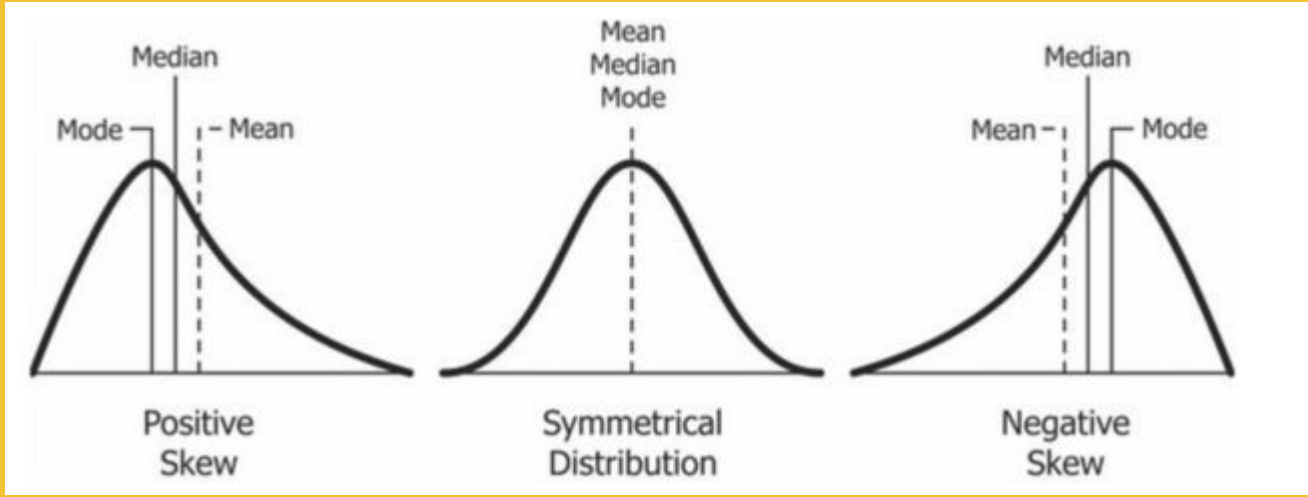


Econ 2250: Stats for Econ

Fall 2022

Source for pic stats above.

- **Today**
  - **Review homework 8**
  - **Review linear model**
  - **Review t-test**
  - **Look at homework 9**

Variance

$V(X) = E((X - E(X))^2)$

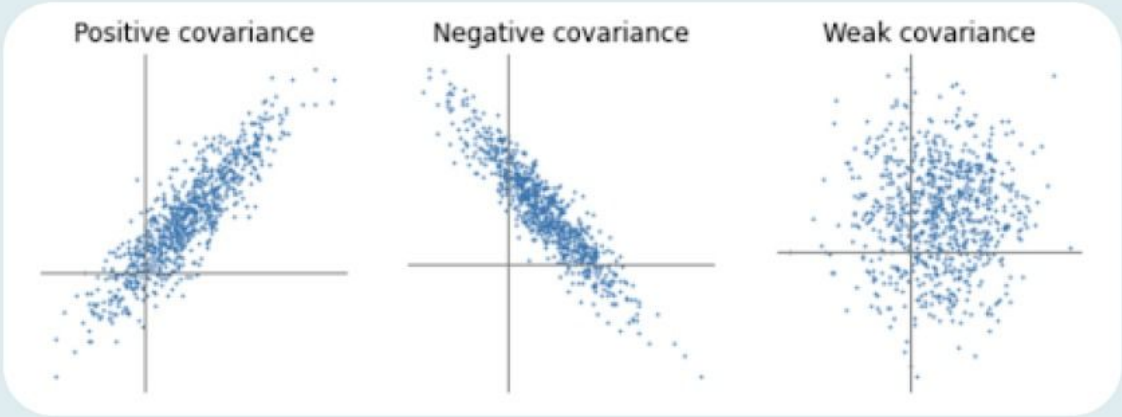
Variance

Theorem. **Computational formula for the variance:**

$$\mathrm{Var}(X) = E[X^2] - (EX)^2.$$

Proof:

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X-\mu_X)^2] \\
&= E[X^2 - 2\mu_X X + \mu_X^2] \\
&= E[X^2] - 2E[\mu_X X] + E[\mu_X^2] \quad \text{by linearity of expectation.} \\
&= E[X^2] - 2\mu_X^2 + \mu_X^2 \\
&= E[X^2] - \mu_X^2.
\end{aligned}$$



Covariance

$Cov(X,Y) = E[(X-E(X)(Y-E(Y)]$

image

Covariance

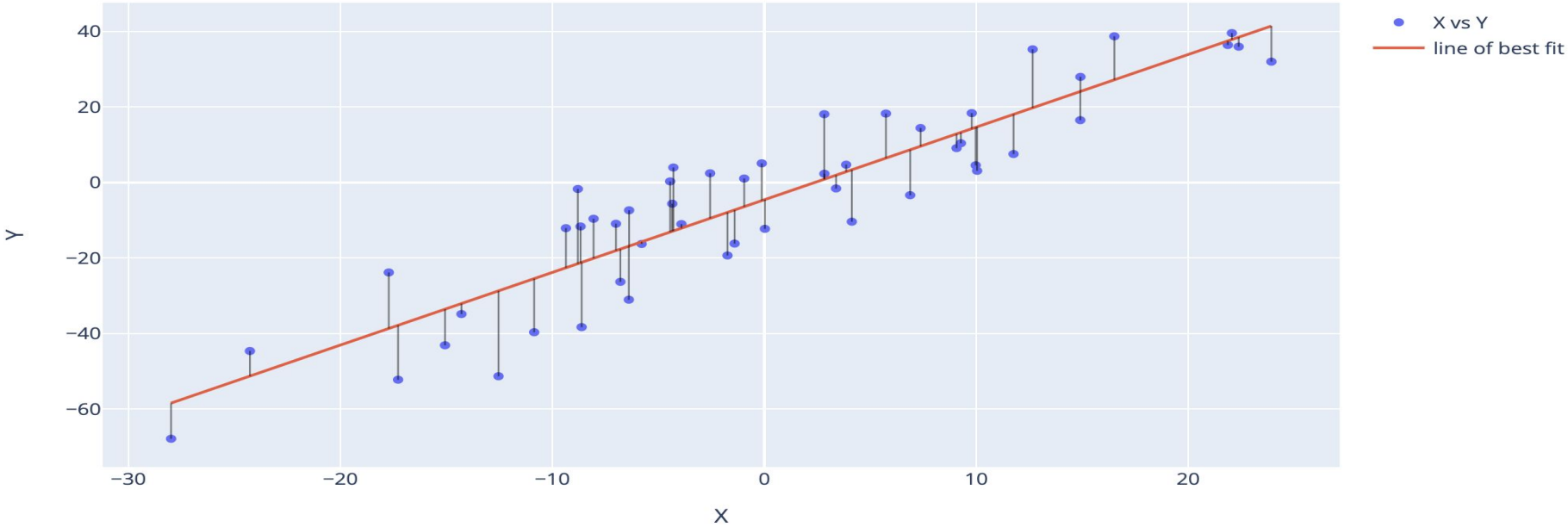
$$Cov(x, y) = E[(X - E(X))(Y - E(Y))]$$
$$= \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{n - 1}$$

Linear Regression

$$y_i = a + b * x_i + u_i$$

$$\hat{y}_i = \hat{a} + \hat{b} * x_i$$

$$\hat{u}_i = y_i - \hat{y}_i$$



$$\hat{y}_i = \text{best guess intercept} + \text{best guess slope} * x_i$$

$$\hat{a} = \text{best guess intercept}$$

$$\hat{b} = \text{best guess slope}$$

$$\hat{y}_i = \hat{a} + \hat{b} * x_i$$

$$\hat{y}_i = \hat{a} + \hat{b} * x_i$$

$$\hat{b} = \frac{\operatorname{cov}(x, y)}{\operatorname{var}(x)} = \frac{\sum(x_i - \bar{x}) \sum(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$\hat{a} = \bar{y} - \hat{b} * \bar{x}$$

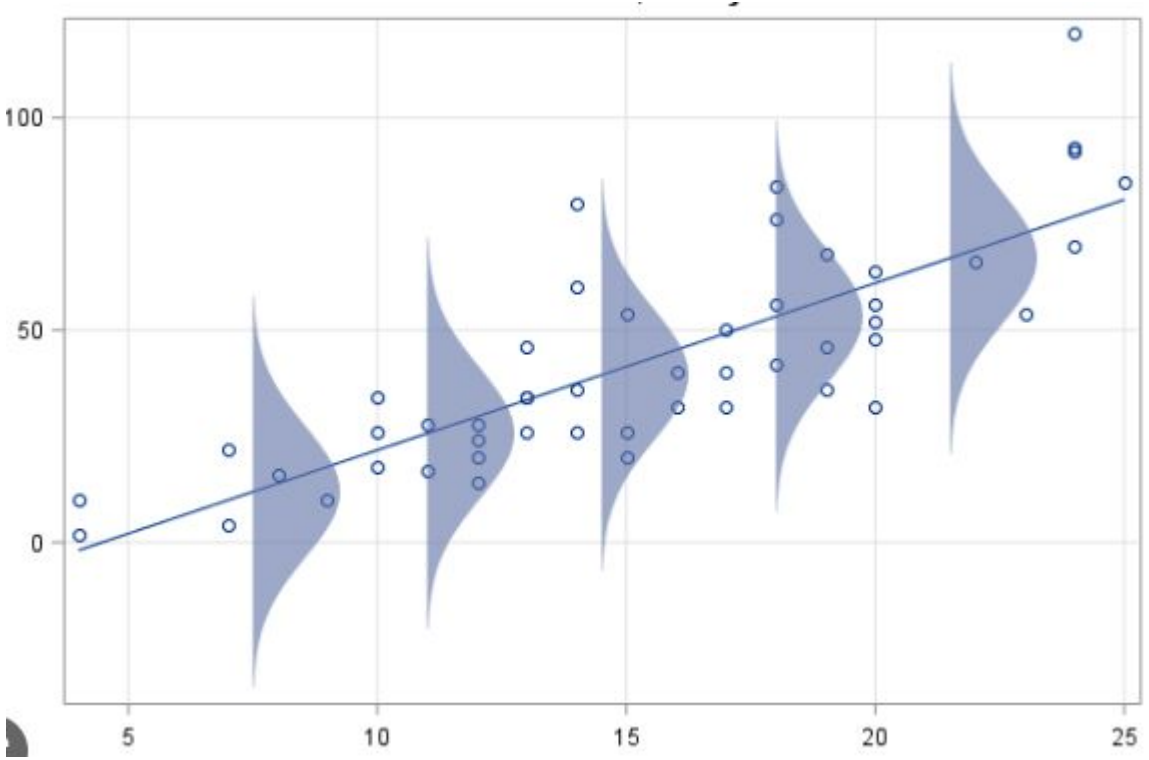
Walk through example

x	y	demean_x	demean_x_sq	demean_y	demean_y_sq	demean_x*demean_y		yhat b0 + b1*x	y-yhat error
3	6	-2.4	5.76	1.6	2.56	-3.84		4.90	1.10
5	9	-0.4	0.16	4.6	21.16	-1.84		4.48	4.52
2	2	-3.4	11.56	-2.4	5.76	8.16		5.11	-3.11
8	1	2.6	6.76	-3.4	11.56	-8.84		3.85	-2.85
9	4	3.6	12.96	-0.4	0.16	-1.44		3.65	0.35
			37.2		41.2	-7.8	sum		
						-1.95	sum/(n-1)		

mean_y	4.4
mean_x	5.4

numerator	-1.95	-1.95	-0.199	correlation
denom	sqrt(37.2/4 * 41.2/4)	9.79		

slope	-0.210
intercept	5.53

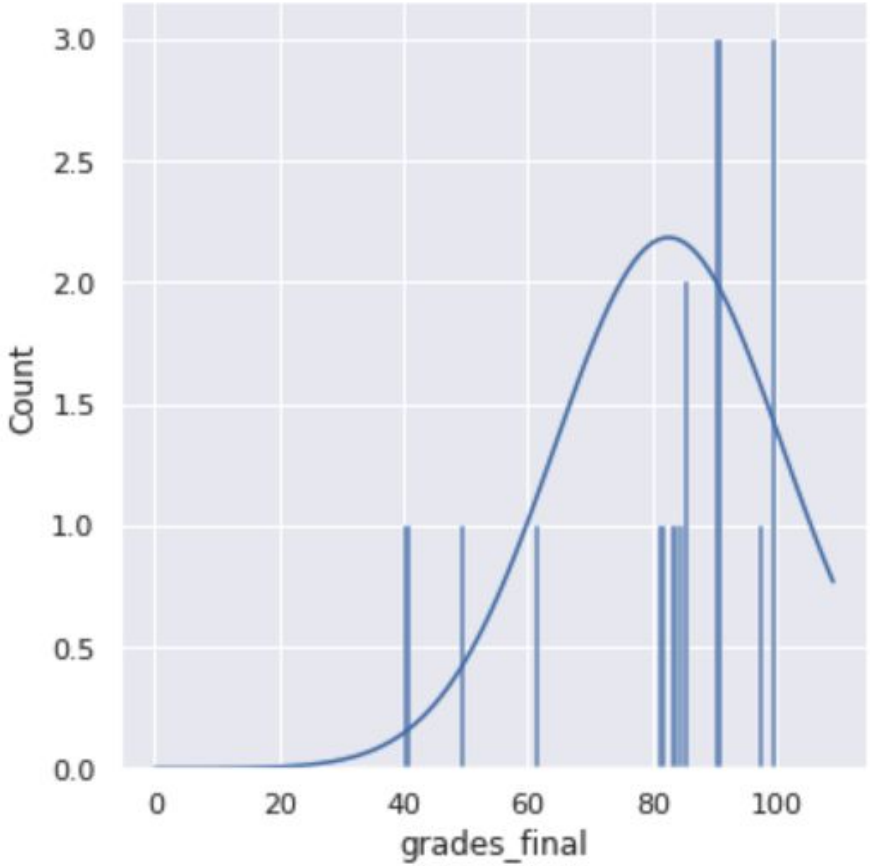

100
50
0
5
10
15
20
25

(simple) T Tests = z-test

$$z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$$

$$\bar{x} = \frac{\sum x}{15} = \frac{1238}{15} = 82$$

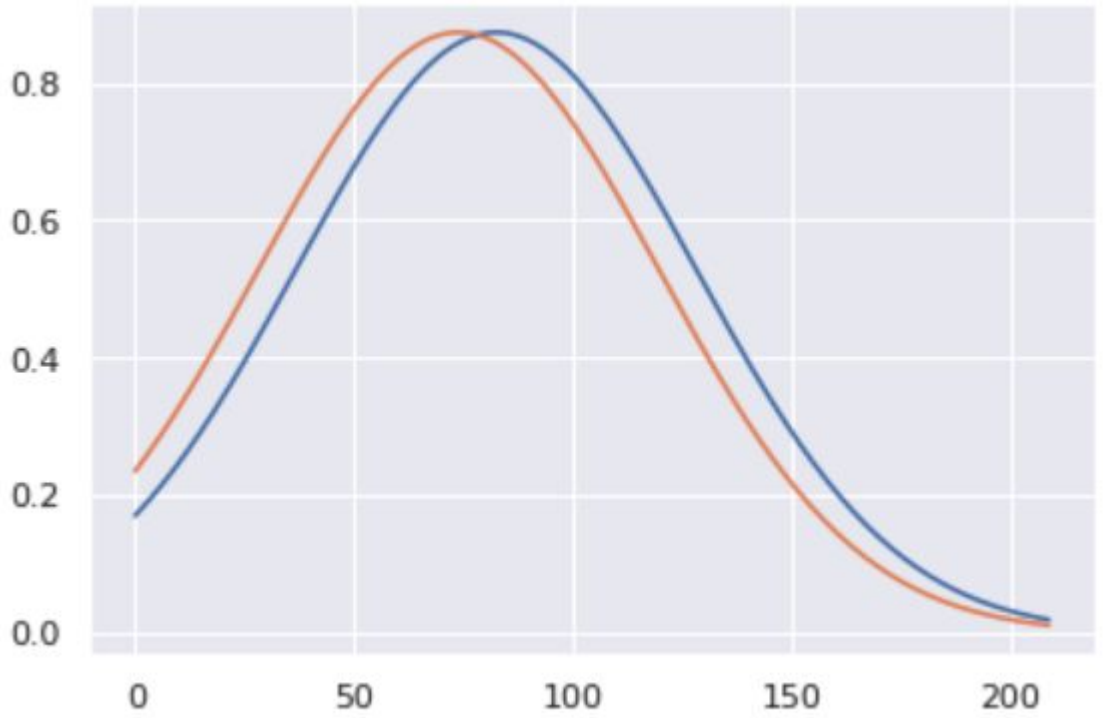
$$\sigma = \sqrt{\frac{(\sum x - \bar{x})^2}{N-1}} = \sqrt{\frac{4672}{14}} = 18.2$$



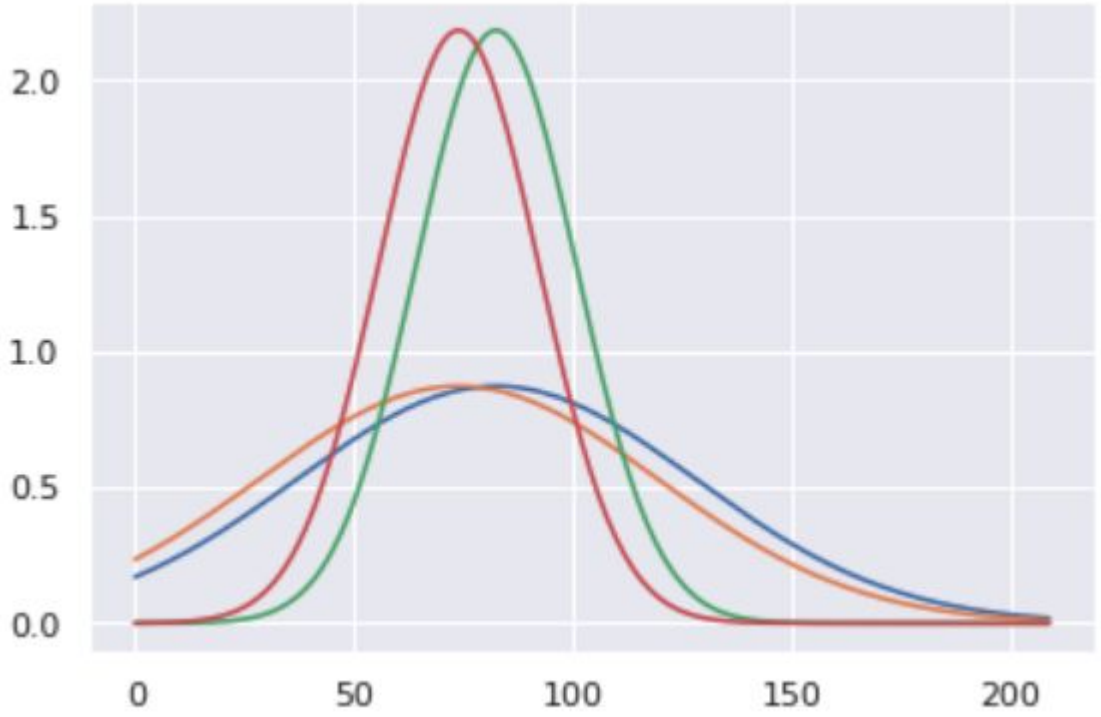
How to tell if significant difference between 2 classes?

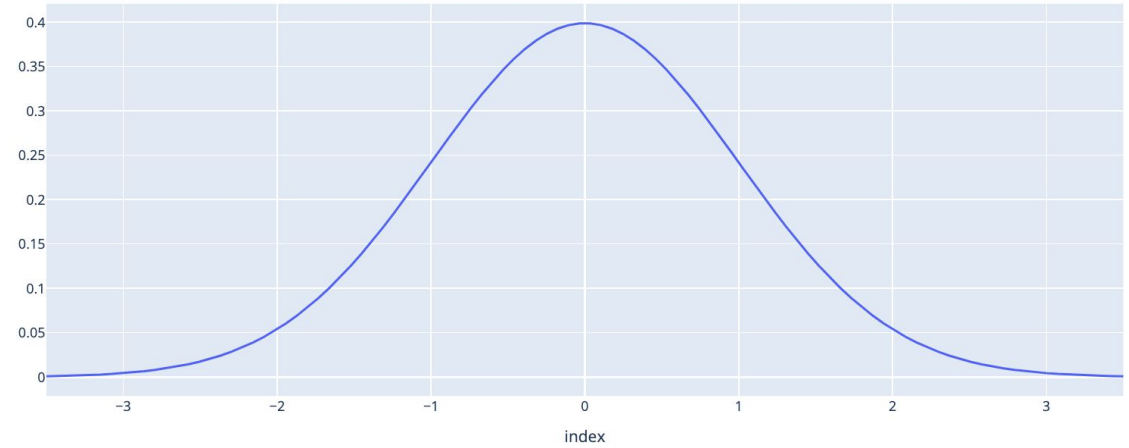
Class1: 82 (18)

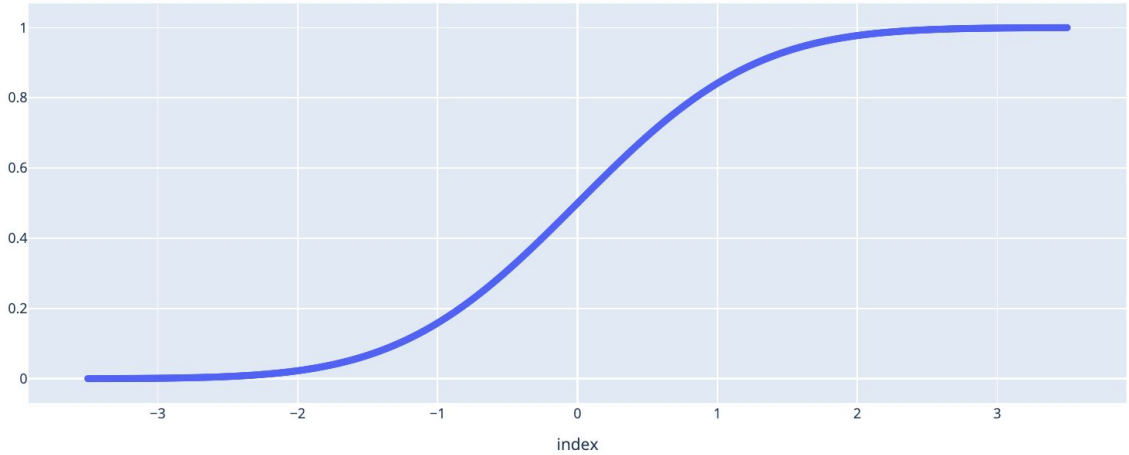
Class2: 72 (18)

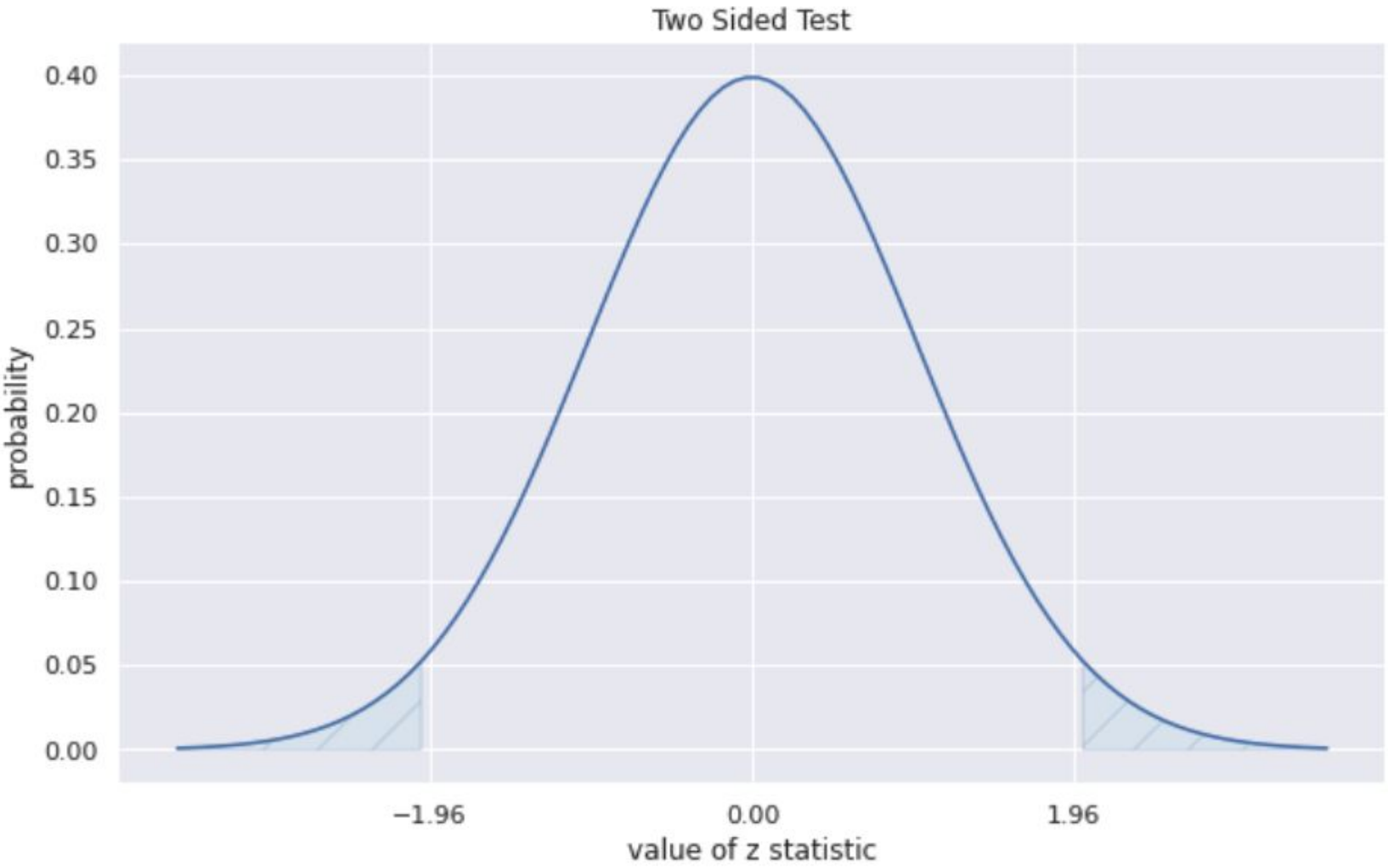


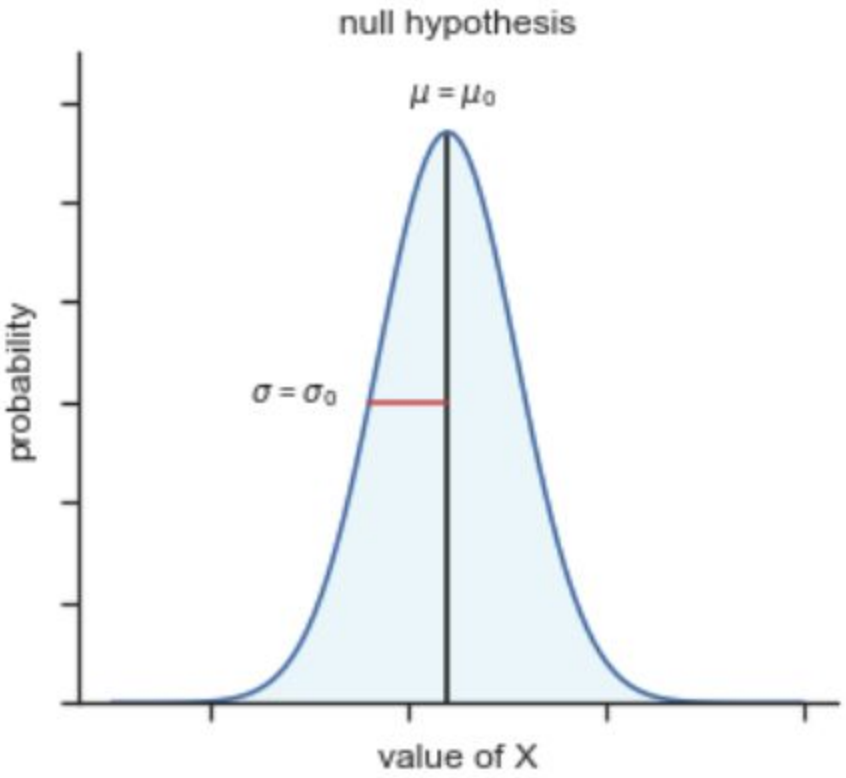
Remember, variance matters

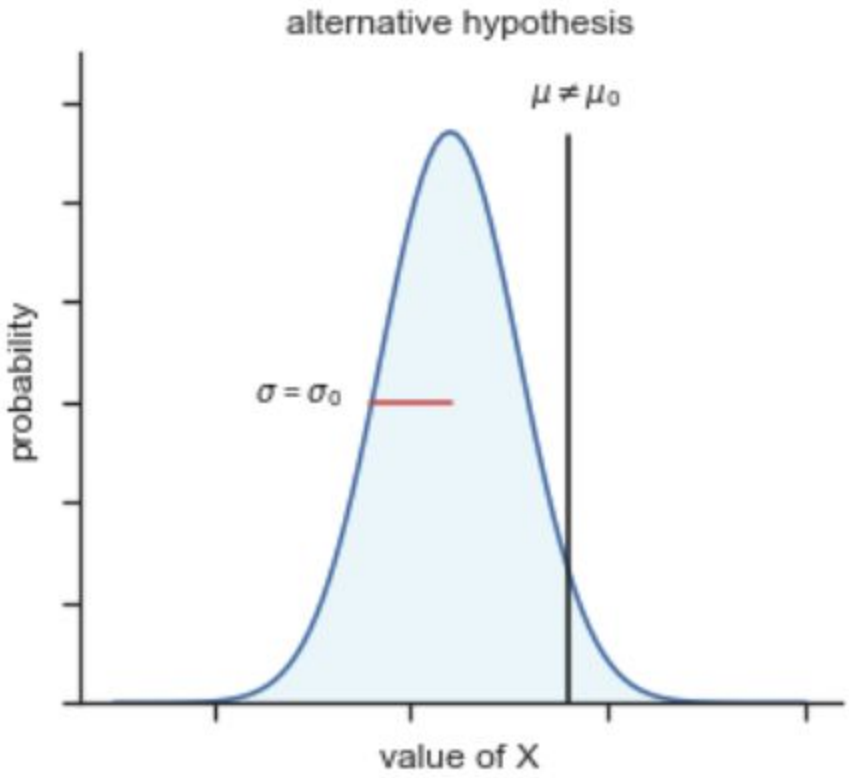



0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
−1
0
1
2
3
index


1
0.8
0.6
0.4
0.2
0
−3
−2
−1
0
1
2
3
index


Two Sided Test
probability
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−1.96
0.00
1.96
value of z statistic


null hypothesis
$\mu = \mu_0$
$\sigma = \sigma_0$
probability
value of X


alternative hypothesis
$\mu \neq \mu_0$
$\sigma = \sigma_0$
probability
value of X

Simple T-Test Case 1

$$z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$$

Class1: 82 (18)

Class2: 72 (18)

H0: Difference - 0

H1: Not H0

$$Z = \frac{82 - 72}{18/\sqrt{15}} = 2.27$$

Simple T-Test Case 2

$$z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$$

Class1: 82 (18)

Class2: 76 (18)

H0: Difference - 0

H1: Not H0

$$Z = \frac{82 - 76}{18/\sqrt{15}} = 1.3$$

Simple T-Test Case 3

$$z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$$

Class1: 82 (28)

Class2: 72 (28)

H0: Difference - 0

H1: Not H0

$$Z = \frac{82 - 72}{28/\sqrt{15}} = 1.4$$

desired alpha level	two-sided test	one-sided test
0.100	1.644854	1.281552
0.050	1.959964	1.644854
0.010	2.575829	2.326348
0.001	3.290527	3.090232