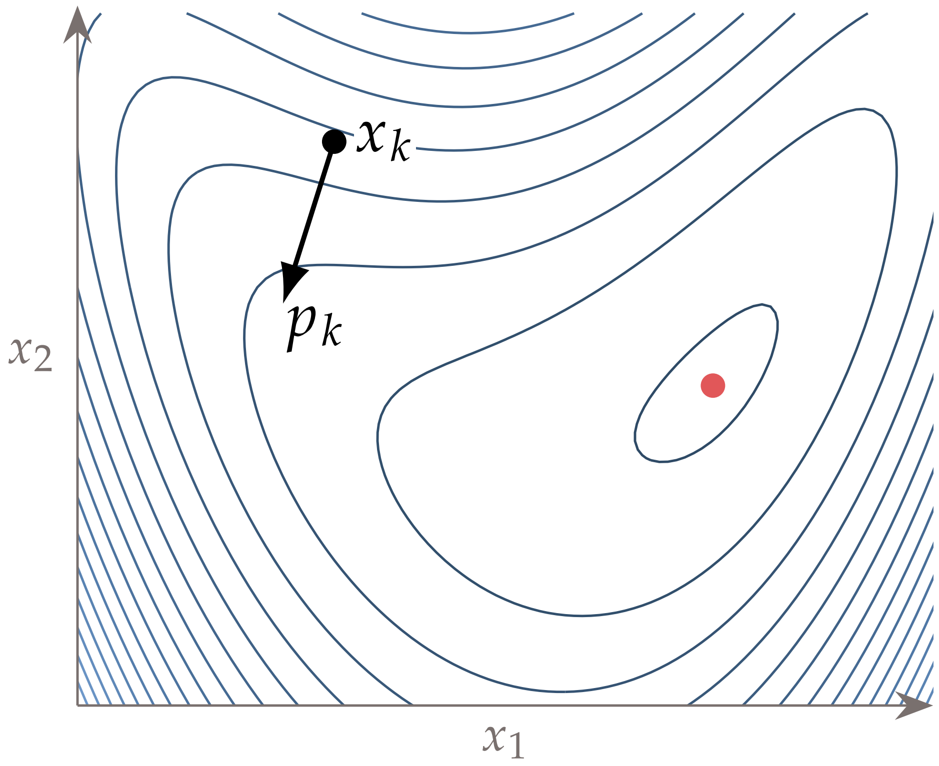
Hypothesis Tests (two samples)

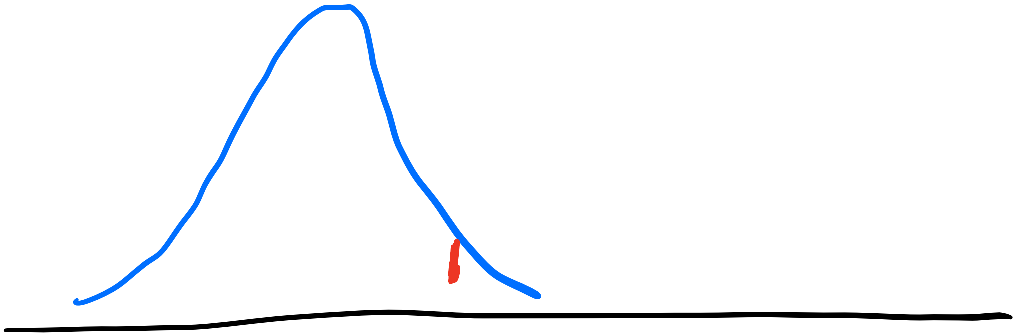


ME EN 275
Andrew Ning
aning@byu.edu

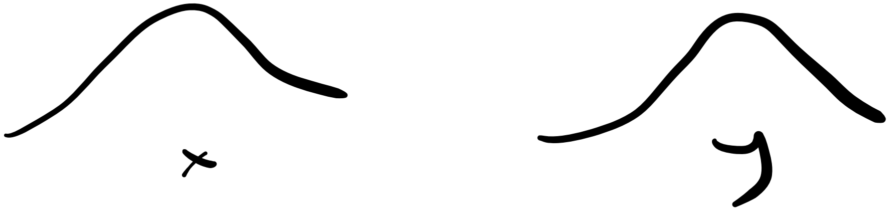
Motivation

I introduce a new process that appears to improves yield from 80.5 to 88.6. Is this a significant change or random variation?

My bicycle frame is subjected to a max load of 100.0 and its strength is 110.0. How likely is it to fail?



Recall: confidence interval for difference between two means



$$(\bar{x} - \bar{y}) \pm z_c \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}$$

Hypothesis test on difference between two means

$$\mu_x - \mu_y = \Delta$$

difference is normally distributed with

mean $= \Delta$

s.d. $= \sqrt{\dfrac{s_x^2}{n_x} + \dfrac{s_y^2}{n_y}}$

Example

diameter of inclusions of two types of welds

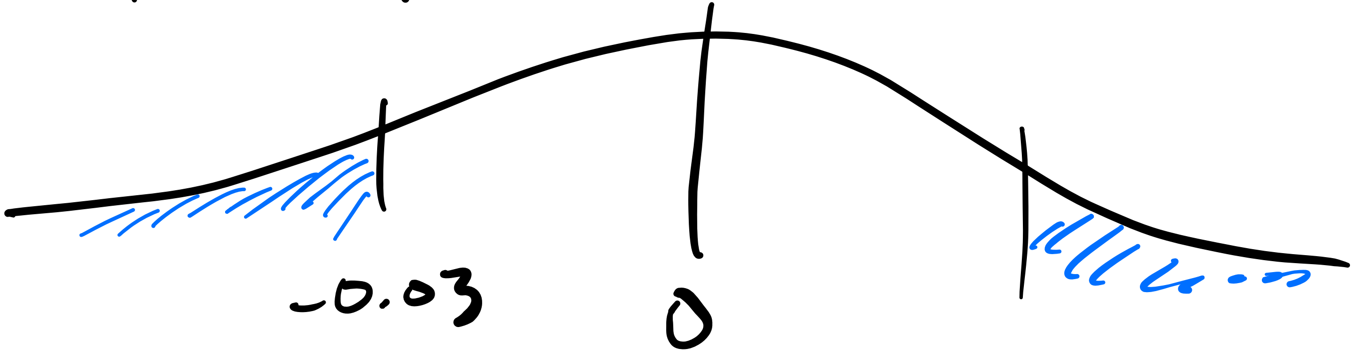
using argon	using CO2
544 welds	581 welds
$\bar{x} = 0.37$	$\bar{y} = 0.4$
$s_x = 0.25$	$s_y = 0.26$

Is the difference in means significant?

$H_0: \mu_x - \mu_y = 0$ $\quad$ $H_a\ \mu_x - \mu_y \neq 0$

$\bar{x} - \bar{y} = -0.03$

$$\sigma = \sqrt{\frac{.25^2}{544} + \frac{.26^2}{581}} = 0.0152$$



2*norm.cdf(-0.03, 0, 0.0152) = 0.0485

95% confidence bound for difference (argon - co2)?

norm.interval(0.95, -0.03, 0.0152)

(-0.0598, -0.000196)

Can we conclude that the mean diameter for carbon dioxide welds (μ_y) exceeds that for argon welds (μ_x) by more than 0.015 μm?

Example

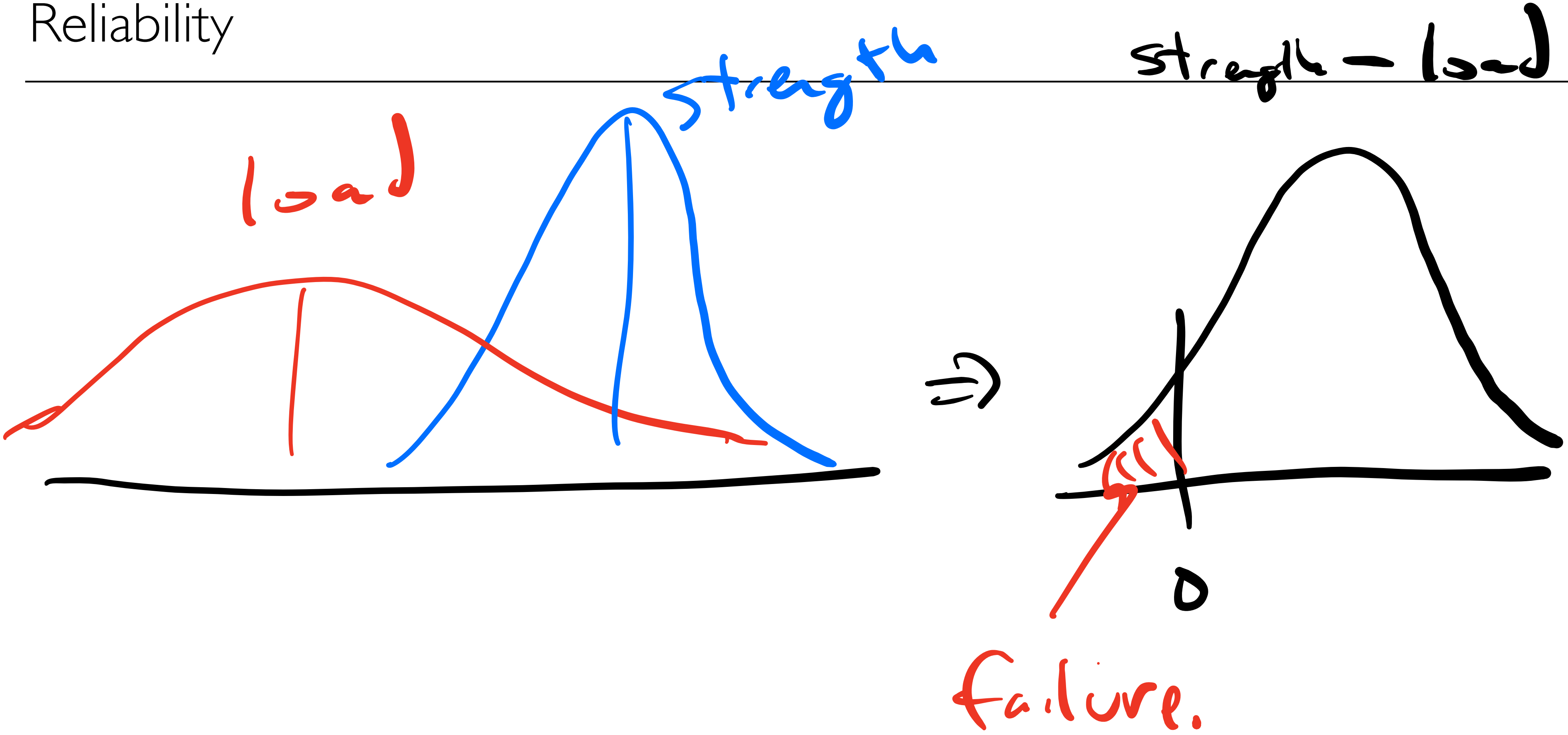
items identified on website

structured design	conventional design
10 users	10 users
$\bar{x} = 44.1$	$\bar{y} = 32.3$
$s_x = 10.09$	$s_y = 8.56$

Is the difference in means significant?

95% lower bound on difference

Reliability



Example

buckling of bicycle component in axial compression

strength	load
16 samples	16 samples
$\bar{x} = 2250$	$\bar{y} = 2000$
$s_x = 200$	$s_y = 800$

probability of failure?

x: strength $\qquad y$: load.

$$\mu_x - \mu_y$$

$$\bar{x} - \bar{y} = 250$$

$$z = \frac{x-\mu}{\sigma}$$

$$\sigma_{\mu_x - \mu_y} = \sqrt{\frac{200^2}{16} + \frac{800^2}{16}} = 206.155$$

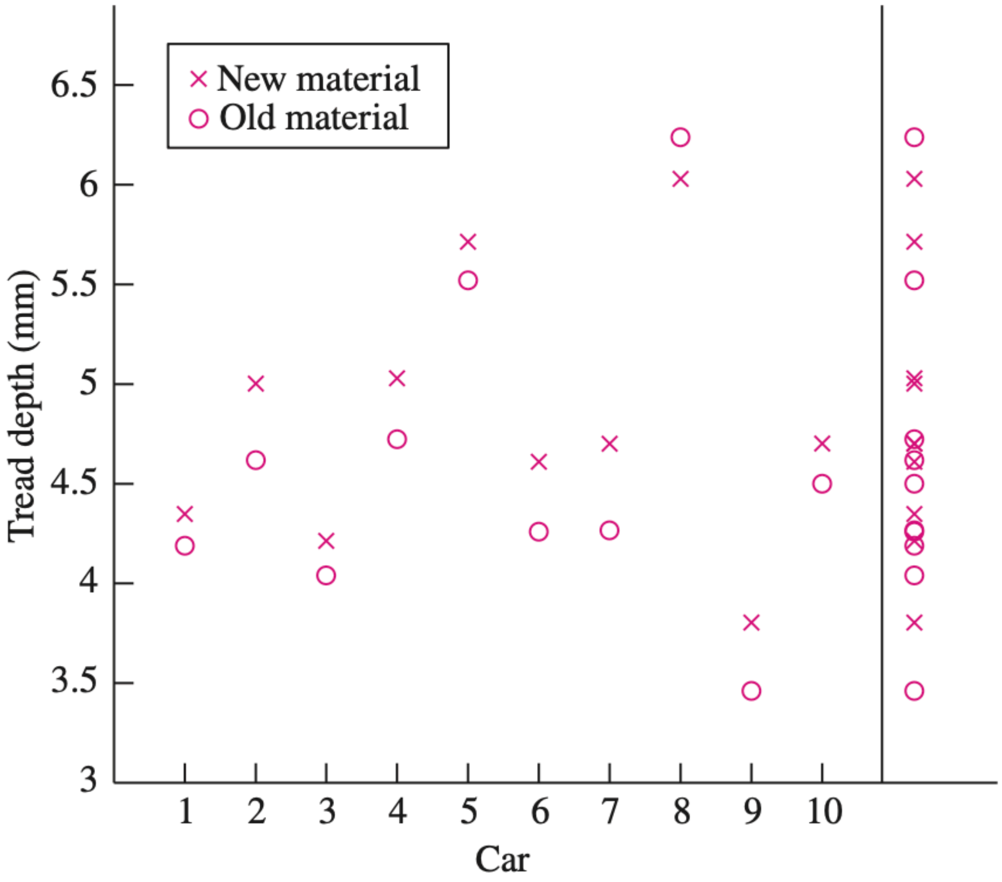
norm.cdf$(0, 250, \sigma)$

$= 0.112$

Paired Data

A tire manufacturer wishes to compare the tread wear of tires made of a new material with that of tires made of a conventional material. One tire of each type is placed on each front wheel of each of 10 front-wheel-drive automobiles. The choice as to which type of tire goes on the right wheel and which goes on the left is made with the flip of a coin.

Title Text



New material
Old material
Tread depth (mm)
6.5
6
5.5
5
4.5
4
3.5
3
1
2
3
4
5
6
7
8
9
10
Car

Title Text

TABLE 7.1 Depths of tread, in mm, for tires made of new and old material

	Car									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
New material	4.35	5.00	4.21	5.03	5.71	4.61	4.70	6.03	3.80	4.70
Old material	4.19	4.62	4.04	4.72	5.52	4.26	4.27	6.24	3.46	4.50
Difference	0.16	0.38	0.17	0.31	0.19	0.35	0.43	−0.21	0.34	0.20

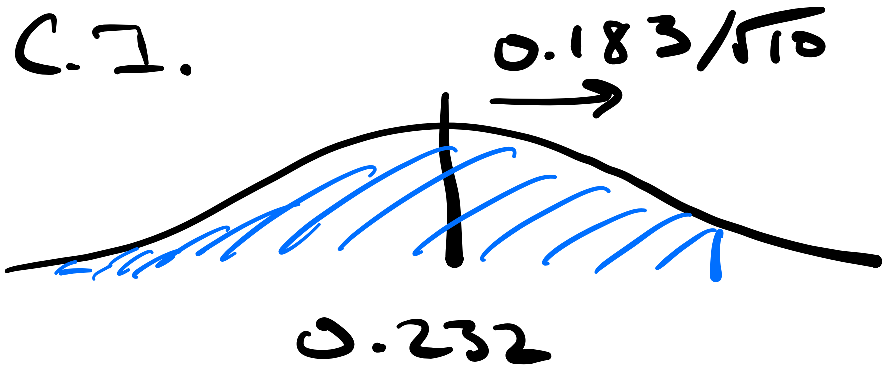
$\bar{x} = 0.232$ ← mean

t dist w/ 9 df

std

$S_x = 0.183$ ← ddof = 1

95% C.I.



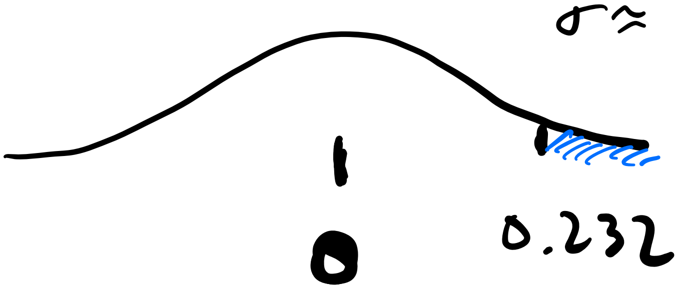
1-sided

$t.ppf(0.95, 9) = 1.83$

$$0.232 - 1.83\frac{(0.183)}{\sqrt{10}} = 0.126$$

$H_0: \mu_D \le 0$ $H_a: \mu_D > 0$

$\sigma \approx \dfrac{0.183}{\sqrt{10}}$



$$1 - t.\mathrm{cdf}\left(0.232,\ \underbrace{9}_{df},\ 0,\ \frac{0.183}{\sqrt{10}}\right)$$

$p = 0.0015$

Recap of Conf. Int. and Hyp. Testing

1 sample group or 2 sample groups.

mean: $\bar{x}$
μ
stdev: $\frac{s_x}{\sqrt{n_x}}$
σ

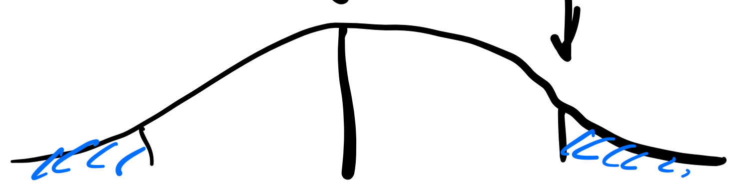
mean: $\bar{x} - \bar{y}$
μ
stdev: $\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}$
σ

Conf. int. or hyp. test.

$\mu \pm z\sigma$

interval or ppf

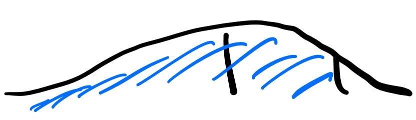
μ_{null} μ



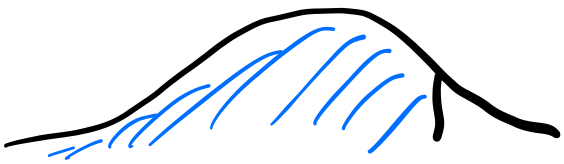
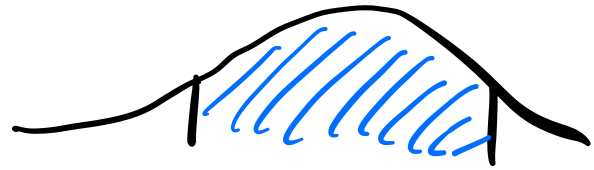
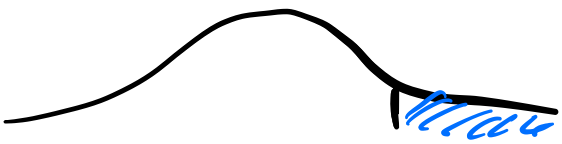
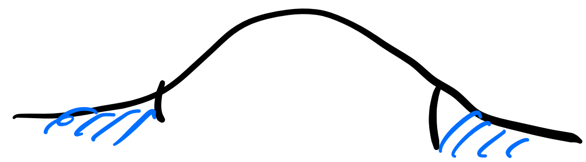
cdf

1-sided or 2-sided

$+$ or $-$



$\pm$

	1 sided	2-sided
C.I.		
H.T.		

large vs small

large	small
$n \geq 30$	$n < 30$
	$df = n-1$
	$df = (n_x - 1) + (n_y - 1)$
↓	↓
normal	t - distribution

$$S_x = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}} \leftarrow \text{ddof} = 1$$

$$\xrightarrow{\text{ddof}=0} \sigma = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n}}$$

independent or <u>paired</u>