~~Ha~~

1 - tailed vs 2 - tailed test.

large values of the z - statistic
are evidence against Ho

Which large values?

Depends on Ho and H_1

Ho : populations A and B have same mean
H_1 : population B's mean is <u>different</u> than population A's mean.

2 tailed test
A > B
and
B > O
count against Ho.

H_1 : population B's mean is <u>greater</u> than population A's mean.

1 tailed test
as only
B > A
cast doubt
on Ho.

Practice Final 1

Q5

a) 99%

$p(A|B) \neq p(B|A)$

$P(\text{disease} \mid \text{positive test})$

$\neq p(\text{positive test} \mid \text{disease})$

"prosecutor's fallacy"

b). population size = 1000 10% have the disease.

(i) 100 ← # people with a cold.

(ii) 99 # of people with a cold who will test positive

(iii) 900

(iv) 9 # of people without a cold who will test positive.

c) $p(\text{disease} \mid \text{positive test})$

→ consider only the members of the population that have had a positive test.

what proportion of them have the disease?

$$\frac{\text{\# positive tests amongst people with colds.}}{\text{total \# of positive tests}}$$

$$= \frac{99}{99+9} = \underline{\underline{0.917}} \qquad (\neq 0.99).$$

What happens if only 0.1% of the population have the disease?

consider population size 100,000

100 have the disease

of whom 99 test positive

99900 don't have the disease.

999 of them will test positive

The vast majority of the people who test positive do not have the disease.

$$p(\text{disease} \mid \text{positive test}) = \frac{99}{999+99} = \frac{99}{1098}$$

$$= \underline{\underline{9\%}}.$$

~~etc~~

Ch 28 Q C 7.

Roll die 600 times

record odd or even.

low (1,2,3) or high (4,5,6)

observed frequencies

	large	small
odd	183	113
even	88	216

is the die fair?

hypothesis test: does this data look like it could come from a fair die?

1 sample z - test

2 sample z - test

χ^2 test for categories with known expected frequencies

χ^2 test for independence.

which type of χ^2 test is it?

is "odd" independent of "large" vs "small"

if the roll is odd, does that tell you something about whether it is large or small?

$$\underbrace{\boxed{1}\quad 2\quad \boxed{3}}_{\text{small}}\quad \underbrace{4\quad \boxed{5}\quad 6}_{\text{large.}} \qquad \boxed{\ }\ \text{odd}$$

If the roll is <u>odd</u> it is twice as likely to be small than large

$\Rightarrow$ <u>NOT</u> independent.

What are the categories.

			600 rolls.
large	odd	$1/6$	100
large	even	$2/6$	200
small	odd	$2/6$	200
small	even	$1/6$	100

<u>χ^2 test.</u>

Practice Final 2 Q2.

Live Search	Mean	14.95	SD	3.29
Yahoo		16.93		3.37

$r = 0.94$

Morris $\rightarrow$ Live search measure of 17.5

17.5 is $\frac{17.5 - 14.95}{3.29}$ SD_x above mean.

prediction is $0.94 \times \frac{17.5 - 14.95}{3.29}$ SD_y above mean y

$16.93 + 3.37 \times 0.94 \times \left(\frac{17.5 - 14.95}{3.29}\right)$

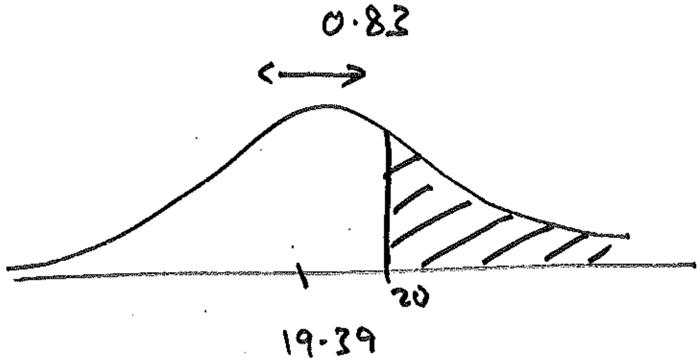
$= 19.39$

RMS error. $= SD_y \times \sqrt{1 - r^2}$

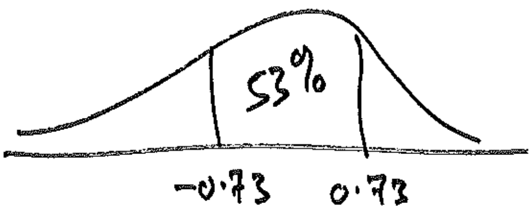
$= 3.37 \times \sqrt{1 - 0.94}$

$= \underline{0.83}$

d)



20 is $\frac{20-19.39}{0.83} = 0.73$ in standard units

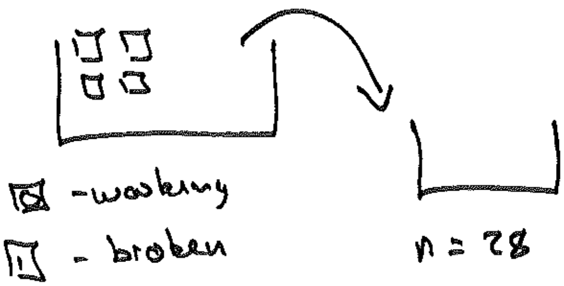


tail area is $\frac{1}{2}(100-53)$.

$= \underline{23.5\%}$.

Practice final 1 Q1.

a). broken 5 days out of 28



$n = 28$

sum of draws $= 5$

95% CI for population %.

$$= \text{sample } \% \pm 2\, SE_{\%}$$

$$\frac{5}{28} \times 100 = \frac{500}{28} \qquad 17.86\%$$

$$SE_{\%} = \frac{SE_{sum}}{n} \times 100$$

$$= \sqrt{\text{sample size}} \times SD_{box} \times \frac{100}{\text{sample size}}$$

$$= \sqrt{28}\ (1-0)\sqrt{\frac{5}{28} \times \frac{23}{28}} \times \frac{100}{28}$$

$$= 7.24\%$$

b) is machine B less reliable than machine A.

n = 28.

8 - # days broken. 28.57 %.

2-sample z test

H_0 : two machines break down at the same rate

H_1 : machine B is broken more often than machine A.

$$z = \frac{\text{observed diff} - \text{expected diff}}{SE_{diff}}$$

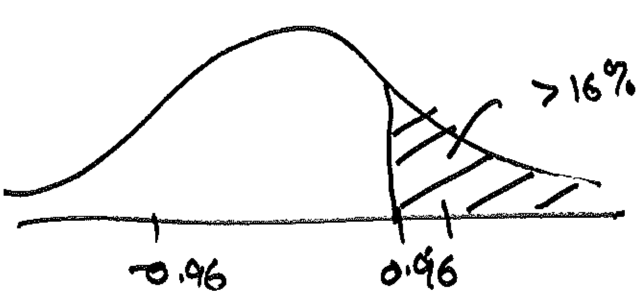
$$= \frac{(28.57 - 17.86) - 0}{\sqrt{7.24^2 + 8.54^2}}$$

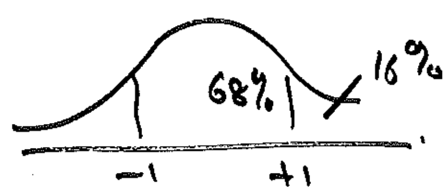
$$= \frac{10.71}{11.20} = \underline{\underline{0.96}}.$$

SE_B.

$$\sqrt{\frac{8}{28} \times \frac{20}{28}} \times \frac{100}{\sqrt{28}}$$

$= 8.54$





Do not have evidence to reject H_0

c) Test for independence.

	B works	B broken
A works	31	13
A broken	14	2.

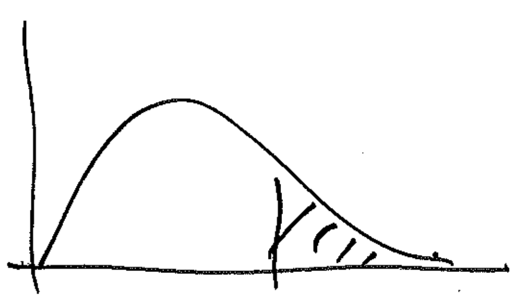
60

χ^2 test for independence.

$$\chi^2 = \text{sum} \frac{(\text{observed freq} - \text{expect. freq})^2}{\text{exp. freq}}$$

\# expect. freq $= \dfrac{(\text{row total}) \times (\text{col total})}{\text{table total}}$

\# dof $= (\#\text{rows} - 1) \times (\#\text{cols} - 1)$.



d). $P(A \text{ broken}) = 5/28$

$P(B \text{ broken}) = 8/28$

assuming independence.

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$P(A \text{ and } B) = P(A) \times P(B)$.

$P(A \text{ or } B) = \frac{5}{28} + 8/28 - 5/28 \times 8/28$.