-- **Begin Chapter 9** -- 1 / 8 / 2008

1 **An Hypothesis** is a statement, or claim, about a population parameter.

===

2 **A Test of Hypothesis about μ** compares $\bar{x}$ with the value claimed to be μ to see if there is a statistically significant difference between $\bar{x}$ and μ. *An hypothesis can be about any parameter, but in this block it will be about μ or p.*

===

3 **The Null Hypothesis H_o** is the statement $\mu = \mu_o$ to be tested. We write H_o: $\mu = \mu_o$. μ_o *is some number*.

===

4 **The Alternative Hypothesis H_1** is a statement that contradicts H_o. $H_1: \mu \neq \mu_o$ or $H_1: \mu < \mu_o$ or $H_1: \mu > \mu_o$.

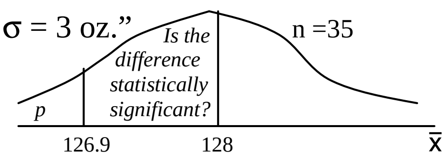
Each alternative corresponds to a tail area of a z or t curve.

$H_1: \mu \neq \mu_o$ involves a two tail area. When μ is suspected to the different from μ_o but we do not know if μ_o will be smaller or larger, then this alternative is used

$H_1: \mu < \mu_o$ involves a left tail area. ← Use this when μ is suspected to be less than μ_o.

$H_1: \mu > \mu_o$ involves a right tail area. ← Use this when μ is suspected to be greater than μ_o.

Example: A milk company claims: "Our 1-gallon jugs contain 128 oz. with σ = 3 oz."
A consumer group suspects that the company is under-filling the jugs.



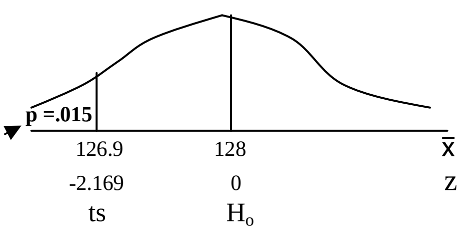
A random sample of 35 jugs has $\bar{x}$ = 126.9 oz. →

126.9 < 128, but is this a statistically significant difference? *We test to see if it is.*

H_o: μ = 128 ← For the hypothesis test, H_o is assumed to be true even though we suspect that H_o is false.

H_1: μ < 128. ← Can this be accepted as true?

H_o: The jugs contain 128 oz. of milk.
H_1: The jugs contain less than 128 oz of milk.



z(126.9) = -2.169 ← This is called The Test Statistic ts.

p = P(z < -2.169) = P($\bar{x}$ < 129.6) = .015. p is called The p-value of the test.

The conclusion of a test of hypothesis will be
1. Reject H_o and accept H_1.
or
2. Fail to reject H_o.

The conclusion depends on p.
If p is small enough, then we reject H_o.

===

5 **The P-value p** or Observed Level of Significance for the test statistic ts_1 is *page 487*

$$p = P(\text{computing a ts more contradictory to } H_o \text{ than } ts_1 \mid H_o \text{ is true}).$$

The smaller the P-value, the larger the disagreement between ts and H_o and agreement between ts and H_1.

If $p \leq .05$, then the difference is statistically significant.

If $p > .05$, then the difference between ts and H_o is not considered to be statistically significant.

very highly significant	highly significant	significant	not statistically significant	p
.001	.01	.05		

===

---

6 | The Level of Significance α | is a number that *may be specified* for the hypothesis test.

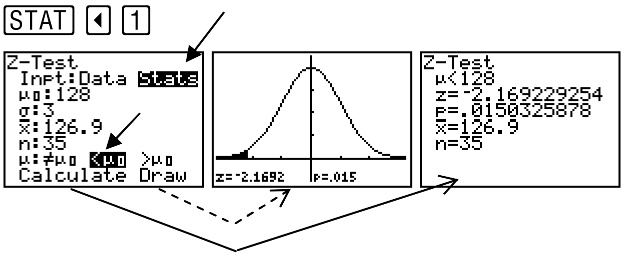
p must be smaller then α. Typically α = .05 or α = .01 — *α is the Greek letter alpha.*

α = P(rejecting $H_o \mid H_o$ is true). α is the probability of making A Type I Error. *See note 7.*

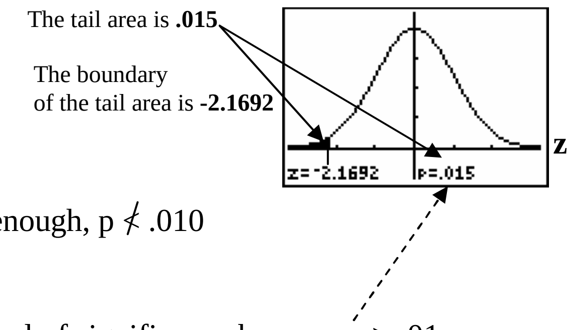
The choice of α depends on the seriousness of a Type I Error.
If a Type I Error is serious, then α is set to 1% or less.

For the milk jug example on sheet 1, let α = 1%.

H_o: μ = 128. H_1: μ < 128. σ = 3 ← a Z-Test. n = 35, $\bar{x}$ = 126.9



→ p is not small enough, p ≮ .010



Conclusion: There is **not** enough evidence to reject H_o at the 1% level of significance because p > .01.

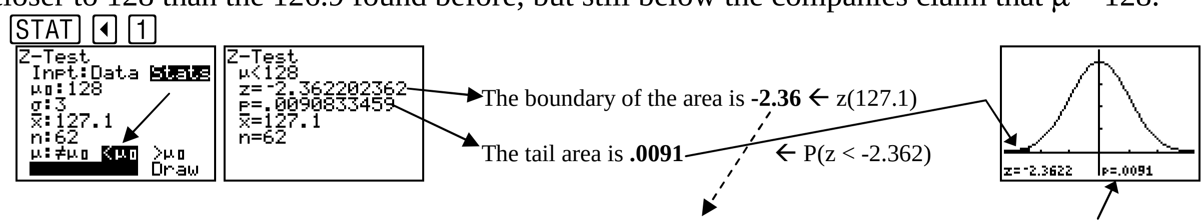
According to the test, the under-filling of the 1-gallon jugs is not statistically significant.

We cannot reject H_o. The consumer group has failed to make a case against the company.

Perhaps their sample n = 35 was too small. We will return to this situation with a larger sample.

The consumer group (*from sheet 1*) repeats the test with a larger sample, n = 62. They now find that $\bar{x}$ = 127.1.

127.1 is closer to 128 than the 126.9 found before, but still below the companies claim that μ = 128.



The test shows that 127.1 has a z-score (test statistic) of z = **-2.3622** and a P-value of p = **.0091**.

Since this p < .01, the consumer group now has enough evidence against H_o: μ = 128. We reject H_o : μ = 128 and accept **H_1: μ < 128.** Their findings now are statistically significant at the 1% level of significance.

The consumer group can say, H_1: μ < 128, and the 1-gallon jugs are significantly under-filled by the company.

Sometimes the evidence leads us to reject H_o even though H_o is true. This rejection is a Type I error.

---

7 | A Type I Error | is rejecting H_o when H_o is true. Rejecting the truth is a Type I error. *page 488-489*

α = the probability of making a Type I error. α = P(rejecting $H_o \mid H_o$ is true)

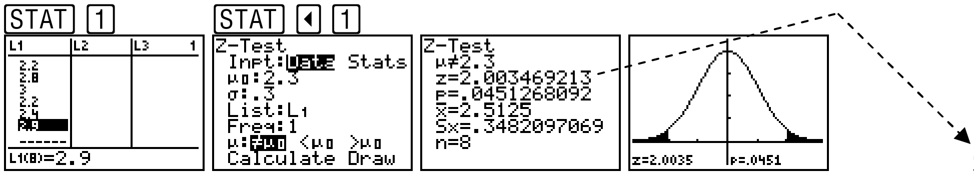
---

8 | A Type II Error | is accepting H_o when H_o is false. Accepting a falsehood is a Type II error. *page 488-489*

β = the probability of making a Type II error. β = P(accepting $H_o \mid H_o$ is false) *β is the Greek letter beta.*

This is an example of <u>a two tail test</u>. The concentration is suspected to have changed, but we do not know which way.

492,3 Ammonia nitrogen concentration is normally distributed with $\mu = 2.3$ and $\sigma = 0.3$.
H_0: $\mu = 2.3$. H_1: $\mu \neq 2.3$ (the concentration has changed). $n = 8$, $\bar{x} = 2.5125$
Is there evidence that $\mu \neq 2.3$ using $\alpha = .01$ as the level of significance?



z(2.5125) = 2.0035

With p = .045 (this is bigger than .010) <u>we cannot reject</u> H_0. We cannot say that the concentration has changed.

For any two tail test, p is <u>the sum of the two tail areas</u>. p = P(z < -2.0035 <u>or</u> p > 2.0025) = .0451

---

(9) The Test Statistic ts is the <u>z-score</u> or <u>t-score</u> of $\bar{x}$ used to determine the p-value of the hypothesis test.

9.2 ---

10 The Test Statistics formulas for t and z are as follows:

If [$n \geq 30$ *(a large sample)* or $n \leq 29$ and the population is "normal"] and

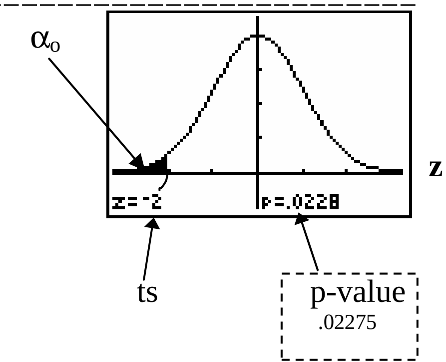
- if σ is <u>known</u>, use $z(\bar{x}) = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ for a ZTEST.
- if σ is <u>unknown</u>, use $t(\bar{x}) = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ for a TTEST.

If $n \leq 29$ *(a small sample)* and the population is <u>not</u> "normal", then <u>neither</u> test statistic is appropriate.

Note: *The test statistics, z <u>and</u> t, can be found using* PRGM **ZTXVALUE**. *They also appear of the test screens.*

---

11 The P-value Test of Hypothesis about μ is done as follows:



a) <u>Decide</u> on the test type: <u>1 tail</u> or <u>2 tail</u>, and is it a <u>TTEST</u> or <u>ZTEST</u>.

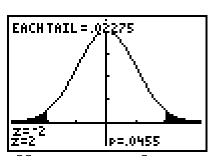
b) <u>Select</u> a sample of size n. Find $\bar{x}$ and, if σ is unknown, find s.

c) <u>Compute</u> ts. $z(\bar{x})$ or $t(\bar{x})$

d) <u>Draw</u> a curve showing the tail area(s) determined by ts.

e) <u>Compute</u> the size of α_o, the size of a tail.

1. For a 1 tail test, α_o is the P-value.
2. For a 2 tail test, $2\alpha_o$ is the P-value.



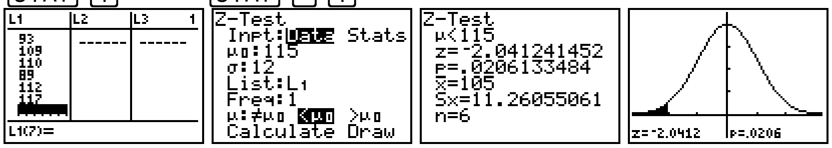
← In this two tail z test, <u>each tail</u> is .02275
z = -2 α_o = .02275 and p = .0455

f) Let α = a level of significance we are willing to tolerate as a Type I error.

g) <u>State</u> the conclusion: 1. If $p \leq \alpha$, then we <u>reject</u> H_o and accept H_1.
2. If $p > \alpha$, then we <u>do not reject</u> H_o.

486,2 We suspect that Rosie's rate <u>slowing</u>. Heart rates are normally distributed. H_0: $\mu = 115$ b/m. $\sigma = 12$.
We find that with n = 6, $\bar{x} = 105$ ← Rosie's average heart rate. Let H_0: $\mu = 115$ and H_1: $\mu < 115$.

STAT 1 STAT ◂ 1

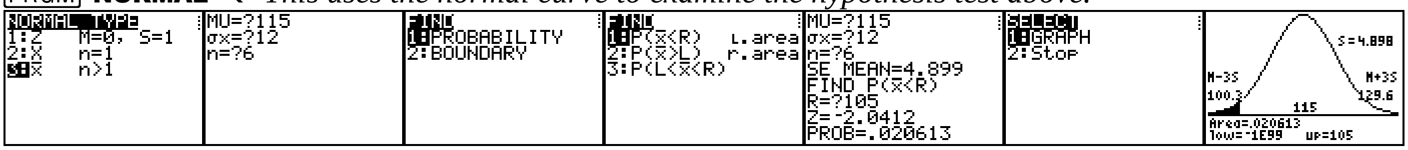


The tail area is **.0206** = p ← P(z < -2.0412)

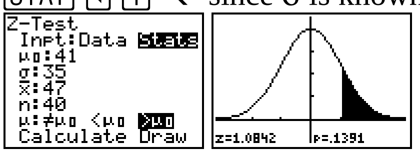
The boundary of the tail area is **-2.0412** = ts = z(105).

This show that there is only a 2% chance of having a rate of 105 or lower (i.e. $P(\bar{x} < 105) \approx 2\%$).
Conclusion: <u>Reject</u> that her rate is normal (115) and <u>accept</u> that Rosie's heart rate is slowing ($\mu < 115$).
The difference between 115 and 105 is statistically significant.

PRGM **NORMAL** ← *This uses the normal curve to examine the hypothesis test above.*



499, 3 STAT ◂ 1 ← since σ is known.



$$z = \frac{47-41}{\frac{35}{\sqrt{40}}} \approx 1.0842$$

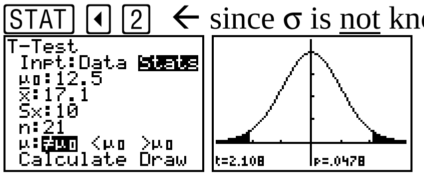
α = .05

P(z > 1.0842) ≈ .1391.

$P(\bar{x} > 47) \approx .1391$.

H_0: $\mu = 41$. $\sigma = 35$. H_1: $\mu > 41$. n = 40, $\bar{x} = 47$
p = .1391 > 0.05 → there is <u>not</u> enough evidence to say that the number of sunspots has increased above 41.

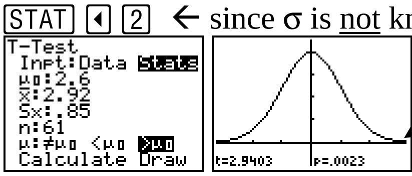
502, 4 STAT ◂ 2 ← since σ is <u>not</u> known. *Since n < 30, the x distribution <u>must be normal</u> to continue.*



← *we get the exact p-value using the calculator.*

H_0: $\mu = 12.5$. s = 10. H_1: $\mu \neq 12.5$. n = 21, $\bar{x} = 17.1$ α = .01
p = .0478 > 0.01 → there is <u>not</u> enough evidence to say that the remission time is different from 12.5 weeks.

503, 4 STAT ◂ 2 ← since σ is <u>not</u> known. *Since n > 29, the x distribution <u>does not have to be normal</u>.*



The area is too small to see.

← *we get the exact p-value.*

H_0: $\mu = 2.6$. s = .85. H_1: $\mu > 2.6$. n = 61, $\bar{x} = 2.92$ α = .01

p = .0023 < 0.01→ there is <u>is</u> enough evidence to reject μ = 2.6 and accept μ > 2.6 (the length is longer than 2.6 cm).

t by formula → $t = t(2.92) = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{2.92-2.6}{\frac{.85}{\sqrt{61}}} \approx 2.9403$. *From the screen above*→ P(t > 2.9403) ≈ .0023.

The (Traditional) Classical Hypothesis Test, uses the Critical Value **cv**,

The Rejection Region **RR**, or (Critical Region) and The Level of Significance **α**. *See notes 12 - 14 below.*

---

12 The Rejection Region RR is the tail region, of the t or z curve, where the rejection of H_o occurs.

RR is also called The Critical Region. *The RR may be a one tail region or a two tail region.*

---

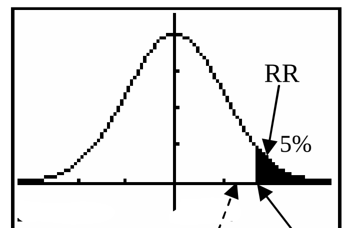
13 The Level of Significance α is the area, or size, of the rejection region.

---

14 The Critical Value cv is the number, t_o or z_o, that is the boundary of the rejection region(s).
cv is the boundary of RR. ts is the boundary of the P-value tail region.

---

15 The Classical Test of Hypothesis about μ is done as follows:



2. ts = 1.6 ∉ RR. cv = 1.645

As an example we show ts and cv Fail to reject Ho.

a) Pick α. *This level of significance is usually 5% or 1%.*

b) Decide on the test type: 1 tail or 2 tail, and TTEST or ZTEST.

c) Select a sample of size n. Find $\bar{x}$, and if σ is unknown, find s.

d) Find cv. cv is in the t-table. *Or use* [PRGM] **CV** .

e) Compute ts. *Use a formula in note 10, or use* [PRGM] **ZTXVALUE**

→ ts is found using [STAT] [◄] [1] or [2]

f) Draw a curve showing the location of both cv and ts.

g) State the conclusion:

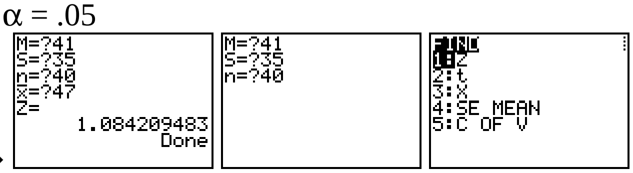
1. If |ts| > |cv|, ts ∈ RR, reject H_o and accept H_1. *H_o is rejected at the α level of significance.*

2. If |ts| < |cv|, ts ∉ RR, do not reject H_o. *The test does not provide enough evidence to reject H_o at the α level of significance. In this case, we have not rejected* H_o, *but we generally do not accept* H_o *unless we consider* β. β *is the probability of making* a Type II *error. (See note8)*

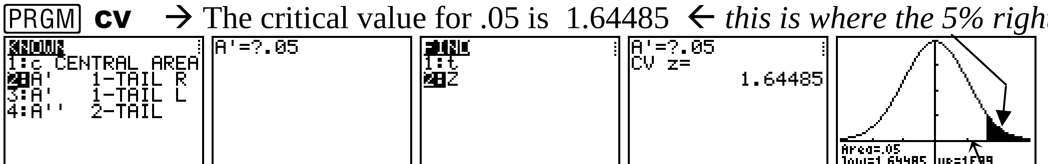
507, 5 H_0: μ = 41. σ = 35. H_1: μ > 41. n = 40, $\bar{x}$ = 47. Use α = .05

$$z = \frac{47 - 41}{\frac{35}{\sqrt{40}}} \approx 1.0842$$

Or use [PRGM] **ZTXVALUE** →



[PRGM] **CV** → The critical value for .05 is 1.64485 ← *this is where the 5% right tail* RR *starts.*



Because z = 1.0842 < 1.64485, z is not in the RR. We cannot reject H_0. 1.0842 ∉ RR

The cv's for the z curve are also found at the bottom of the t distribution table.

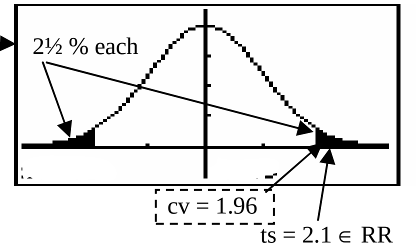
This is the (Traditional) Classical Hypothesis Test. *See 499, 3 (on sheet 4) is the same problem using the P-Value Test.*

Here is the z curve with a two tail RR of 5%. α = 5%.

ts *is given* as 2.1. Notice that 2.1 ∈ RR, so we reject H_o.

The book uses α'' for two tail regions. PRGM **cv** uses A′′ for two tails.

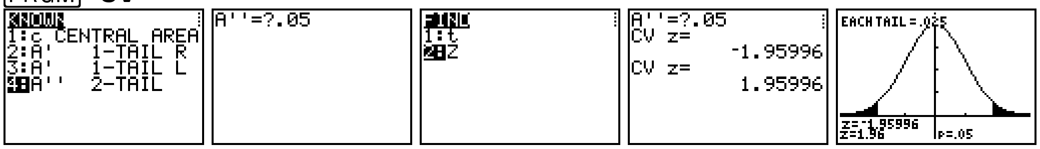
The critical value of the 5% two tail region is 1.96. cv = 1.96.



Note that the total area is 5% and each tail = 2½ %.

As an example we show a ts in RR.

PRGM **cv**

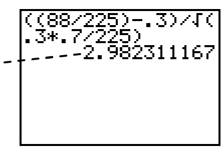


← on the bottom of the t-table.

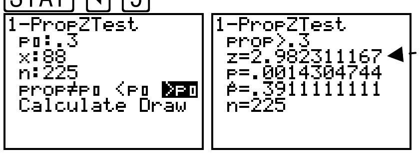
9.3

517, 6 H_0: p_0 = .30. H_1: p_0 > .30. n = 225, r = 88 → $\hat{p} = \frac{88}{225} \approx .39$. $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{(88/225)-.3}{\sqrt{\frac{.3\times .7}{225}}} \approx 2.98$

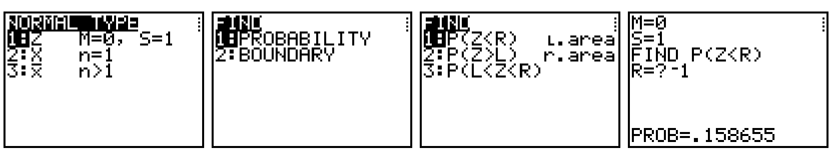
p = P(z > 2.95) ≈ .0016



STAT ◄ 5



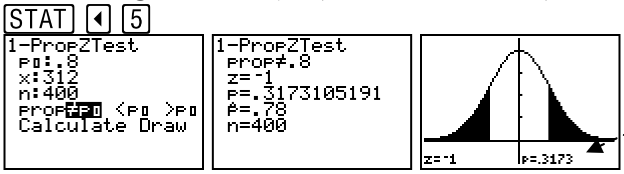
519, 5 H_0: p_0 = .80. H_1: p_0 ≠ .80. n = 400, r = 312 → $\hat{p} = \frac{312}{400} \approx .78$. $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{.78-.8}{\sqrt{\frac{.8\times .2}{400}}} \approx -1.00$



2(.158655) = .31731 = p the two tail area.

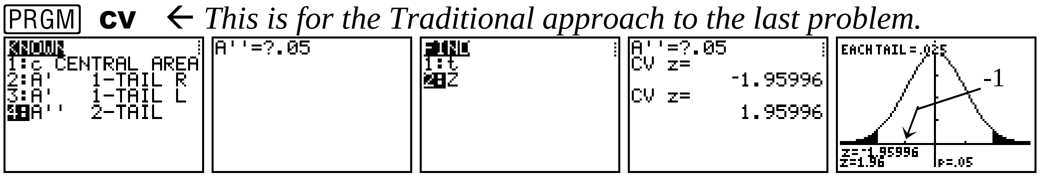
Below we get both z(.78) = -1 and P(z < -1 or z > 1) = .3172

STAT ◄ 5



← p is not small enough to reject H_0. p > .05. Fail to reject.

520, 7 PRGM **cv** ← *This is for the Traditional approach to the last problem.*

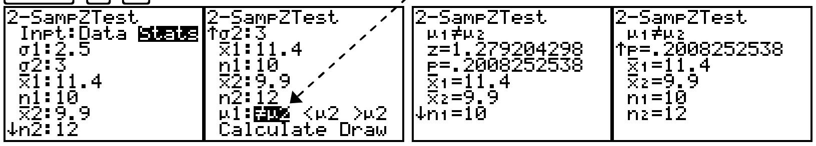


The RR's begin at -1.96 and 1.96. Since -1 is not in either region, we cannot reject H_0.

9.5 ---

In this example we have H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$. Note that ($\mu_1 \neq \mu_2$) is equivalent to saying ($\mu_1 - \mu_2 \neq 0$).

544, 13 STAT ◄ 3 Use $\alpha = .05$.

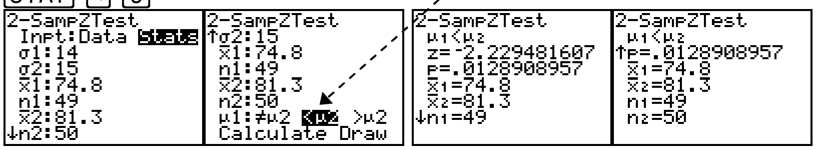


$p = .20 > .05$ → Do no reject H_0.

<u>Do not reject</u> H_0: $\mu_1 = \mu_2$ There is not enough evidence to say $\mu_1 \neq \mu_2$.

In this example we have H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$. Note that ($\mu_1 < \mu_2$) → ($\mu_1 - \mu_2 < 0$).

545, 9 STAT ◄ 3 Use $\alpha = .05$.

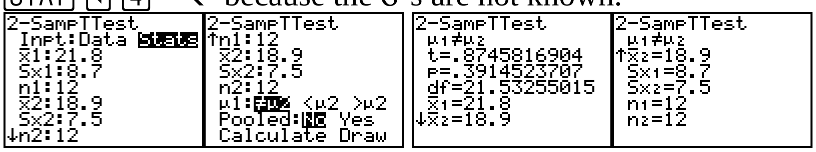


$p = .01 < .05$→ <u>Reject</u> H_0, and accept H_1

The 1st method had lower results than the 2nd. The 2nd method is significantly better than the 1st.

In this example we have H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$.

547, 14 STAT ◄ 4 ← because the σ's are not known. Use $\alpha = .05$.

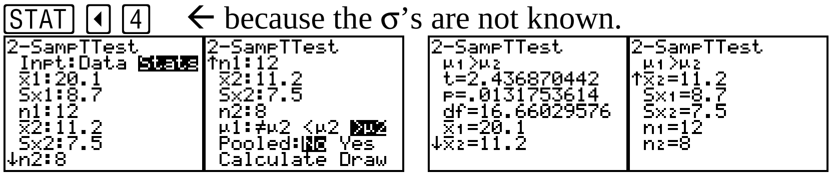


$p = .39 > .05$ → <u>Do not reject</u> H_0.

We cannot say that the means are different.

In this example we have H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$. ← this says that Brand B takes less time to work.

549, 10 STAT ◄ 4 ← because the σ's are not known. Use $\alpha = .05$.

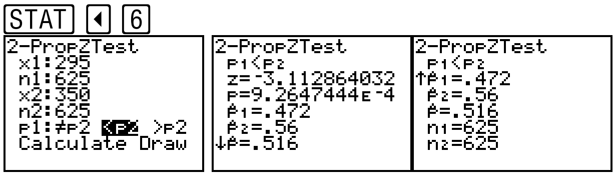


$p = .01 < .05$ → <u>Reject</u> H_0.

Accept H_1: $\mu_1 > \mu_2$. Brand B is significantly faster (takes less time) than Brand A

In this example we test H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. ← this says that Group 1 has fewer people that register.

553, 15 STAT ◄ 6 Use $\alpha = .05$.

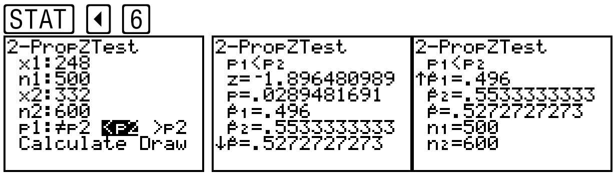


$p = 9.29 \times 10^{-4} = .000929 < .05$ → <u>Reject</u> H_0.

We accept that H_1: $p_1 < p_2$, that Group 1 has fewer people that register.

In this example we test H_0: $p_1 = p_2$. H_1: $p_1 < p_2$. ← this says that Group 1 has fewer people that register.

554, 11 STAT ◄ 6 Use $\alpha = .01$



$p \approx .03 > .01$. → <u>Do not reject</u> H_0.

We cannot say that the reminder improved the registration rate.

== **Begin Chapter 10** ==

1 Paired Data consists of data arranged in pairs X and Y.

Here is a random sample of paired data from 10 students in a statistics class.

Student number	1	2	3	4	5	6	7	8	9	10
Study Time in hours **X**	**7**	**2**	**6**	**3**	**7**	**2**	**4**	**6**	**4**	**5**
Test Grade in % **Y**	**94**	**60**	**87**	**68**	**96**	**65**	**72**	**80**	**68**	**80**

a) What, if any, is the relationship between X and Y? *Linear.* *See note 2*
b) How strong is the relationship between X and Y? *r = .96.* *See note 6*
c) How much of the variation in Y is due to the variation in X? $r^2 = 93\%$ *See note 7*
d) Can we predict a test grade $\hat{Y}$ from a study time X? *4.5h → 76%* *See note 5.*

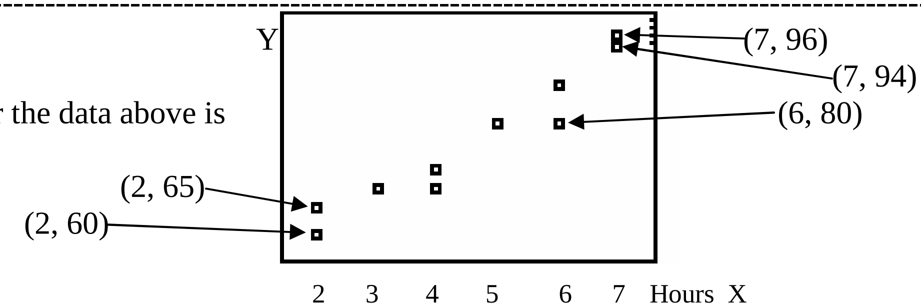
==

2 Correlation Analysis is the study of the relationship between two variables X and Y.

==

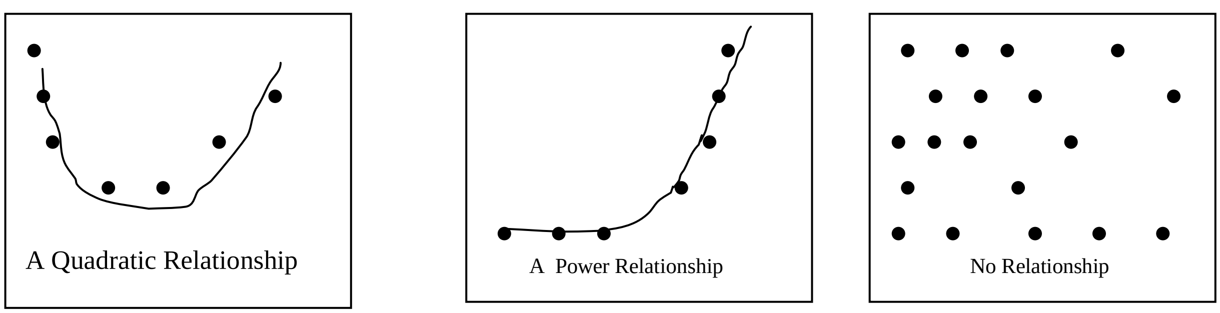
3 A Scatter Diagram, or scatter plot is the graph of paired data. Page 582

The scatter diagram for the data above is



==

In this course we look for a linear relationship between X and Y but many other relationships are possible.



==

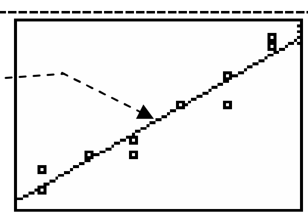
4 The Least Squares Line is the line that *comes closest* to the points of the scatter diagram.

Other names for this line are The Regression Line , The Best Fit Line , or The Line of Best Fit .

The least squares line for the (time, grade) study is shown here.

The equation of this line is $\hat{y} = 48.0 + 6.3x$. From Algebra: y = mx + b

It can be used to make predictions about Y based on X. Pages 604 - 605.



==

5 Regression Analysis involves making predictions about Y based on X.

==

(6) The Pearson Product-Moment ***Coefficient of Correlation*** r , $-1 \le r \le 1$, is the number that indicates the strength of the linear relationship between X and Y. The closer r is to 1 or –1 the stronger the relationship. When r is close to 0, there is little or no linear relationship between X and Y. Page 587

Pages 589 - 593 show how to compute r by *formula* ←*Optional*. *We use* PRGM **LINREG**

For |r| , the linear correlation is given below.

.90 ≤	$\lvert r\rvert \le 1$	very strong.
.70 ≤	$\lvert r\rvert < .90$	strong.
.50 ≤	$\lvert r\rvert < .70$	moderate.
.20 ≤	$\lvert r\rvert < .50$	weak.
0 ≤	$\lvert r\rvert < .20$	nonexistent.

nonexistent	weak	moderate	strong	very strong
0 – .20	.20 – .50	.50 – .70	.70 – .90	.90 – 1

10.2

7 The Coefficient of Determination r^2 is the percent of variation in Y that is due to the variation in X. Page 611

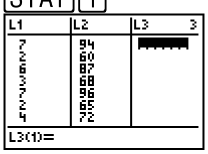
For the (time, grade) study, $r^2 = 93\%$.

This indicates that 93% of the variation in test grades Y is due to the variation in study time X.

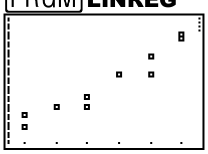
7% of the variation is due to other causes that we cannot explain.
Some other factors that could affect test results are sickness, lack of sleep, or mental attitude.

For a linear regression problem, *put the X data in* L1. *Put the Y data in* L2. *Use the data from Sheet 8.*

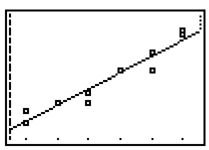
STAT 1



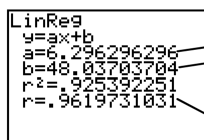
PRGM **LINREG**



scatter plot



regression line



$\hat{y} = 6.3x + 48.0$

r^2 = the coefficient of determination

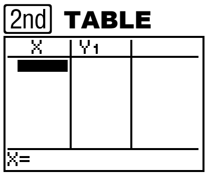
r = the coefficient of correlation

This screen gives a and b that form **the regression equation**: $\hat{y} = 48.0 + 6.3x$.

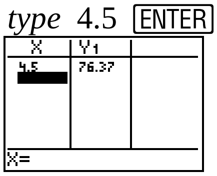
Substitute a time X to predict a grade Y. The prediction is good for $2 \le X \le 7$. This is called interpolation. Page 607

To find the predicted grade Y with 4.5 hours X of studying, $\hat{y} = 48.0 + 6.3(4.5) = 76.35$

2nd **TABLE**



type 4.5 ENTER

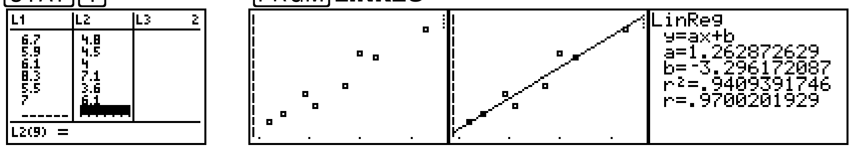


Study time of 4.5 hours should produce a grade of about 76%.
Using an X outside the range of [2, 7] is called extrapolation. Page 607

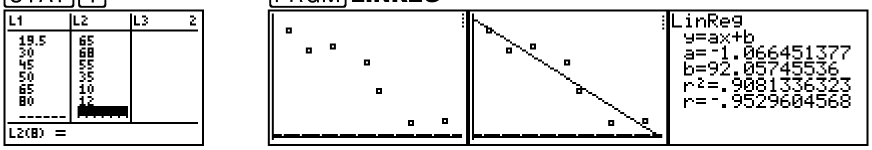
For example, using 15 hours to predict a grade is extrapolation and may give unreliable results.

The coefficient of correlation, r = .96, gives the strength of the relationship between hours studying and the grade. r = .96 indicates that there is a very strong relationship between grades and studying.

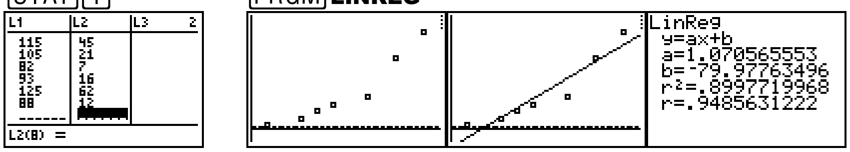
582, 1 STAT 1 PRGM **LINREG**



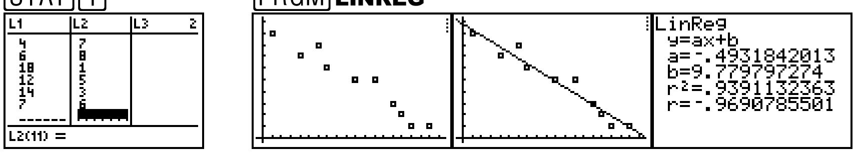
584, 1 STAT 1 PRGM **LINREG**



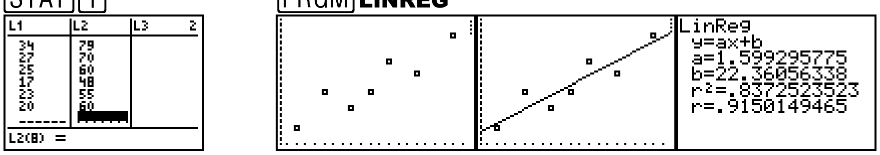
590, 2 STAT 1 PRGM **LINREG**



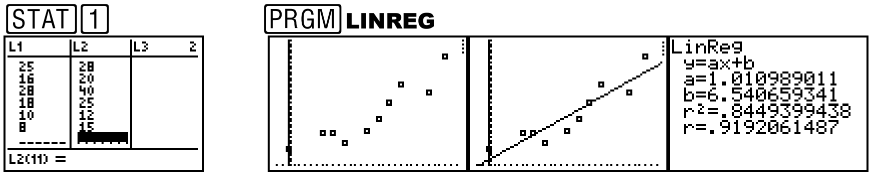
592, 3 STAT 1 PRGM **LINREG**



605, 4 STAT 1 PRGM **LINREG**



608, 4 & 613, 5 STAT 1 PRGM **LINREG**



To compute the regression line, $\hat{y} = mx + b$, and r, the coefficient of correlation , by hand, use the formulas:

$$m = \frac{n \times \Sigma(xy) - \Sigma x \times \Sigma y}{n \times \Sigma(x^2) - (\Sigma x)^2}, \quad b = \frac{\Sigma y - m \times \Sigma x}{n}, \quad r = \frac{n \times \Sigma(xy) - \Sigma x \times \Sigma y}{\sqrt{\left(n \times \Sigma(x^2) - (\Sigma x)^2\right)\left(n \times \Sigma(y^2) - (\Sigma y)^2\right)}}$$

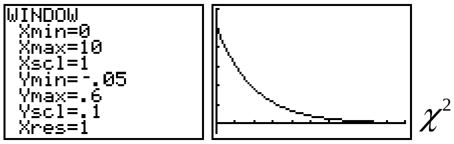
Because of calculators, we no longer need to do the work by hand.

11.1 ==

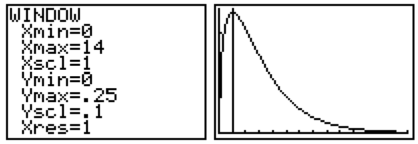
9 The Chi-Square Distribution, denoted χ^2, is a sampling distribution formed by all the values of the form $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ for all samples of size n taken from a normal distribution with known σ.

n is the sample size. df = n – 1. The <u>Mode</u> = df – 2 Note: As df → ∞, χ^2 → ND

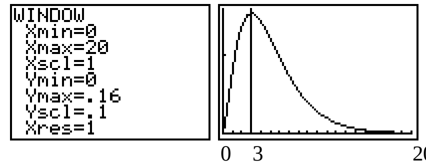
For df = 2



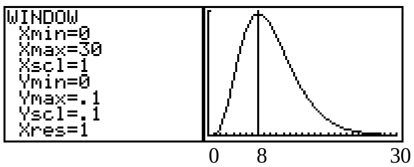
For df = 3



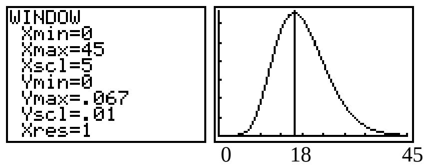
For df = 5



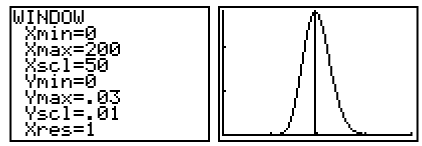
For df = 10



For df = 20

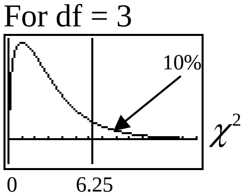


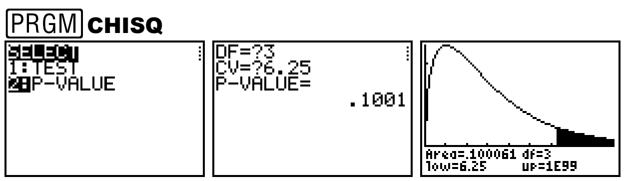
For df = 100



The value on the χ^2 axis is 6.26. This is a critical value. We say χ^2 = 6.25.
The area to the right of 6.25 is the P-value. In this example p = .1001 ≈ 10%.

For df = 3





==

Chi-Square Test of Independence

----

10 A Contingency Table (*two-way frequency table*) shows the frequency of two variables: a row variable and a column variable. The row variable will have a number of categories. So will the column variable.

This example has its row variable as (Crash or no crash) and column variable as (Color of helmet).

The table contains the <u>Observed</u>, O, frequencies.

A 2×3 table	Color of Helmet		
	Black	White	Yellow or Orange
Crash	213	112	8
No Crash	491	377	31

rows × columns table
r × c 2 × 3

==

11 A Test Of Independence tests H_o : The row and column variables are independent.

H_o : *There is no association between the row variable and the column variable.*

==

12 The Expected Frequency E of a cell is $E = \dfrac{\text{(row total)(column total)}}{\text{Sample size}}$

A 2×3 table	Black	White	Y/O	Row total
Crash	213 190.3	112 132.2	8 10.5	333
No Crash	491 513.7	377 356.8	31 28.5	899
Column total	704	489	39	1232

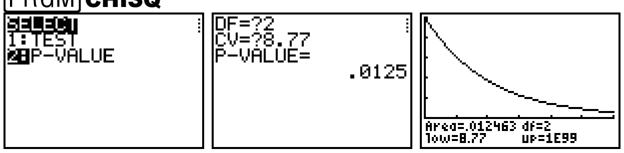
The small numbers are the expected frequencies. $\dfrac{899 \times 39}{1232} = 28.4586$

The number of degrees of freedom df = (r – 1)(c – 1) . df = (2 – 1)(3 – 1) = 2.

The Test Statistic for a Test of Independence is $\chi^2 = \sum \dfrac{(O - E)^2}{E}$

$$\chi^2 = \frac{(213-190.3)^2}{190.3} + \frac{(112-132.2)^2}{132.2} + \frac{(8-10.5)^2}{10.5} + \frac{(491-513.7)^2}{513.7} + \frac{(377-356.8)^2}{356.8} + \frac{(31-28.5)^2}{28.5} = 8.77$$

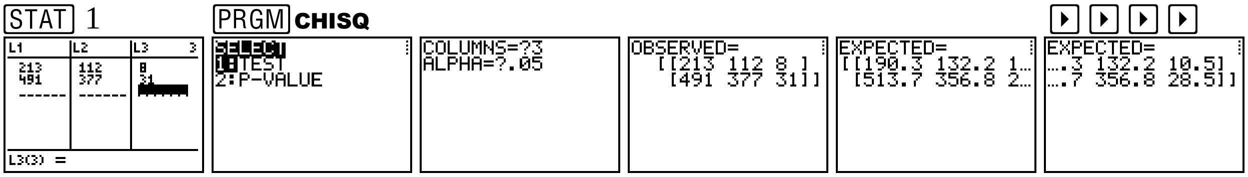
PRGM CHISQ



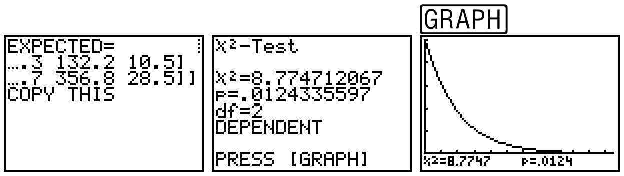
← *The area, 0.012, is too small to see.*

This is how we will do A Test of Independence.

STAT 1 PRGM CHISQ ▸ ▸ ▸ ▸



GRAPH



← *If p < α , then reject H_o → the variables are dependent.*

H_o: There is no association (that is, independence) between Crashes and Helmet color, at the 5% L of S.

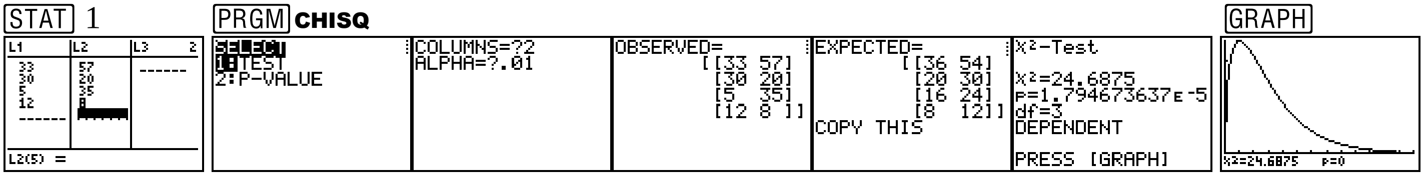
H_1 There is an association (dependence) between Crashes and Helmet color, at the 5% L of S.

With α = .05 and p = .012, we have p < α. Therefore we reject H_o.

Conclusion: Crashes and helmet color are dependent. There is an association between helmet color and crashes.

678, 5

STAT 1 PRGM CHISQ GRAPH



== End Block 4 == End Course == Best Wishes == ☺? ☹? ==

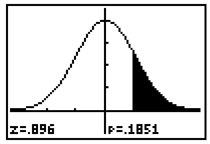
9.1

496, 9 STAT 1 STAT ◄ 1

```
L1    L2    L3   1
4
3.4
6.5
7.1
5.3
6.1
L1(11) =
```

```
Z-Test
 Inpt:Data Stats
 μ0:4.7
 σ:2.4
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ>4.7
 z=.8959786704
 p=.1851320422
 x̄=5.38
 Sx=1.136075115
 n=10
```



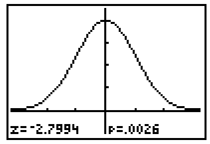
← *more exact than the book*

496, 11 STAT 1 STAT ◄ 1

```
L1    L2    L3   1
3.7   ------ ------
2.9
3.8
4.2
4.8
3.1
L1(7)=
```

```
Z-Test
 Inpt:Data Stats
 μ0:4.55
 σ:.7
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ<4.55
 z=-2.799416849
 p=.0025598103
 x̄=3.75
 Sx=.7007139217
 n=6
```

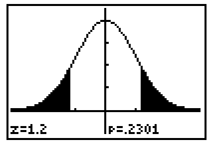


497, 13 STAT 1 STAT ◄ 1

```
L1    L2    L3   1
24
20
13
9
12
5
L1(17) =
```

```
Z-Test
 Inpt:Data Stats
 μ0:11
 σ:5
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ≠11
 z=1.2
 p=.2301394632
 x̄=12.5
 Sx=5.138093031
 n=16
```



9.2

510,1 STAT ◄ 1 $\alpha = .01$

```
Z-Test
 Inpt:Data Stats
 μ0:16.4
 σ:3.5
 x̄:17.3
 n:36
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ>16.4
 z=1.542857143
 p=.0614327356
 x̄=17.3
 n=36
```

511, 3 STAT ◄ 1 $\alpha = .05$

```
Z-Test
 Inpt:Data Stats
 μ0:41.7
 σ:18.5
 x̄:36.2
 n:45
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ≠41.7
 z=-1.994330899
 p=.0461157646
 x̄=36.2
 n=45
```

511, 5 STAT ◄ 2 ← *since σ is <u>not</u> known.* $\alpha = .01$

```
T-Test
 Inpt:Data Stats
 μ0:1.75
 x̄:2.05
 Sx:.82
 n:46
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
T-Test
 μ>1.75
 t=2.481340238
 p=.0084456757
 x̄=2.05
 Sx=.82
 n=46
```

511, 7 STAT ◄ 2 $\alpha = .05$

```
T-Test
 Inpt:Data Stats
 μ0:19.4
 x̄:17.9
 Sx:5.2
 n:36
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
T-Test
 μ≠19.4
 t=-1.730769231
 p=.0923018291
 x̄=17.9
 Sx=5.2
 n=36
```

512, 9 STAT 1 STAT ◄ 2 $\alpha = .05$

```
L1    L2    L3   1
4.9   ------ ------
4.2
4.5
4.1
4.4
4.3
L1(7)=
```

```
T-Test
 Inpt:Data Stats
 μ0:4.8
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
T-Test
 μ<4.8
 t=-3.464101615
 p=.0089814423
 x̄=4.4
 Sx=.2828427125
 n=6
```

513, 13 STAT 1 STAT ◄ 2 $\alpha = .05$

```
L1    L2    L3   1
30    ------ ------
37
43
38
35
36
L1(7)=
```

```
T-Test
 Inpt:Data Stats
 μ0:40
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
T-Test
 μ≠40
 t=-2.026362567
 p=.0985672229
 x̄=36.5
 Sx=4.23083916
 n=6
```

515, 19 STAT ◄ 7 $H_0: \mu = 20$ 20 ∉ (20.28, 23.72) → 20 ∉ 99% CI → Reject H_0

```
ZInterval
 Inpt:Data Stats
 σ:4
 x̄:22
 n:36
 C-Level:.99
 Calculate
```

```
ZInterval
 (20.283,23.717)
 x̄=22
 n=36
```

b STAT ◄ 1 $\alpha = .01$

```
Z-Test
 Inpt:Data Stats
 μ0:20
 σ:4
 x̄:22
 n:36
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

```
Z-Test
 μ≠20
 z=3
 p=.0026999344
 x̄=22
 n=36
```

9.3

522, 1 STAT ◄ 5

```
1-PropZTest
 p0:.301
 x:46
 n:215
 prop≠p0 <p0 >p0
 Calculate Draw
```

```
1-PropZTest
 prop<.301
 z=-2.78258709
 p=.0026964282
 p̂=.2139534884
 n=215
```

523, 3 STAT ◄ 5

```
1-PropZTest
 p0:.7
 x:24
 n:32
 prop≠p0 <p0 >p0
 Calculate Draw
```

```
1-PropZTest
 prop≠.7
 z=.6172133998
 p=.5370938606
 p̂=.75
 n=32
```

523, 5 STAT ◄ 5

```
1-PropZTest
 p0:.77
 x:15
 n:27
 prop≠p0 <p0 >p0
 Calculate Draw
```

```
1-PropZTest
 prop<.77
 z=-2.647813555
 p=.0040507513
 p̂=.5555555556
 n=27
```

523, 7 STAT ◄ 5

```
1-PropZTest
 p0:.5
 x:10
 n:34
 prop≠p0 <p0 >p0
 Calculate Draw
```

```
1-PropZTest
 prop<.5
 z=-2.400980192
 p=.0081756039
 p̂=.2941176471
 n=34
```

524, 13 STAT ◄ 5

```
1-PropZTest
 p0:.092
 x:29
 n:196
 prop≠p0 <p0 >p0
 Calculate Draw
```

```
1-PropZTest
 prop>.092
 z=2.710582113
 p=.0033583114
 p̂=.1479591837
 n=196
```

9.5 ---

558,1 STAT ◄ 3 H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 > \mu_2$.

```
2-SampZTest
 Inpt:Data Stats
 σ1:.5
 σ2:.7
 x̄1:2.8
 n1:10
 x̄2:2.1
↓n2:10
```

```
2-SampZTest
↑σ2:.7
 x̄1:2.8
 n1:10
 x̄2:2.1
 n2:10
 μ1:≠μ2 <μ2 >μ2
 Calculate Draw
```

```
2-SampZTest
 μ1>μ2
 z=2.573251177
 p=.005037432
 x̄1=2.8
 x̄2=2.1
↓n1=10
```

```
2-SampZTest
 μ1>μ2
↑p=.005037432
 x̄1=2.8
 x̄2=2.1
 n1=10
 n2=10
```

p = .005 < .01 → Reject H_0

558, 3 STAT ◄ 3 H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$.

```
2-SampZTest
 Inpt:Data Stats
 σ1:1.5
 σ2:1.2
 x̄1:4.9
 n1:46
 x̄2:4.3
↓n2:51
```

```
2-SampZTest
↑σ2:1.2
 x̄1:4.9
 n1:46
 x̄2:4.3
 n2:51
 μ1:≠μ2 <μ2 >μ2
 Calculate Draw
```

```
2-SampZTest
 μ1≠μ2
 z=2.160170168
 p=.0307593883
 x̄1=4.9
 x̄2=4.3
↓n1=46
```

```
2-SampZTest
 μ1≠μ2
↑p=.0307593883
 x̄1=4.9
 x̄2=4.3
 n1=46
 n2=51
```

p = .03 < .05 → Reject H_0.

559, 5 STAT 1   STAT ◄ 4 H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$.

```
L1      L2      L3     2
 1.8     3.5
 4.8     2.4
 2.9     3.1
 3.1     3.5
 ------  5.2
         2.8

L2(13) =
```

```
2-SampTTest
 Inpt:Data Stats
 List1:L1
 List2:L2
 Freq1:1
 Freq2:1
 μ1:≠μ2 <μ2 >μ2
↓Pooled:No Yes
```

```
2-SampTTest
↑List1:L1
 List2:L2
 Freq1:1
 Freq2:1
 μ1:≠μ2 <μ2 >μ2
 Pooled:No Yes
 Calculate Draw
```

```
2-SampTTest
 μ1<μ2
 t=-.9524109541
 p=.1761393863
 df=19.96330195
 x̄1=3.51
↓x̄2=3.866666667
```

```
2-SampTTest
 μ1<μ2
↑x̄2=3.866666667
 Sx1=.813019065
 Sx2=.943237494
 n1=10
 n2=12
```

p = .176 > .01. Do not reject H_0.

563, 17 STAT ◄ 6 H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$.

```
2-PropZTest
 x1:59
 n1:220
 x2:56
 n2:175
 p1:≠p2 <p2 >p2
 Calculate Draw
```

```
2-PropZTest
 p1≠p2
 z=-1.126116047
 p=.2601164974
 p̂1=.2681818182
 p̂2=.32
↓p̂=.2911392405
```

```
2-PropZTest
 p1≠p2
↑p̂1=.2681818182
 p̂2=.32
 p̂=.2911392405
 n1=220
 n2=175
```

p = .26 > .05 → Do not reject H_0.

564, 21 STAT ◄ 6 H_0: $p_1 = p_2$. H_1: $p_1 < p_2$.

```
2-PropZTest
 x1:37
 n1:100
 x2:47
 n2:100
 p1:≠p2 <p2 >p2
 Calculate Draw
```

```
2-PropZTest
 p1<p2
 z=-1.432670638
 p=.0759760371
 p̂1=.37
 p̂2=.47
↓p̂=.42
```

```
2-PropZTest
 p1<p2
↑p̂1=.37
 p̂2=.47
 p̂=.42
 n1=100
 n2=100
```

p = .08 > .01 → Do not reject H_0.

564, 23 STAT ◄ 6 H_0: $p_1 = p_2$. H_1: $p_1 < p_2$.

```
2-PropZTest
 x1:194
 n1:378
 x2:320
 n2:516
 p1:≠p2 <p2 >p2
 Calculate Draw
```

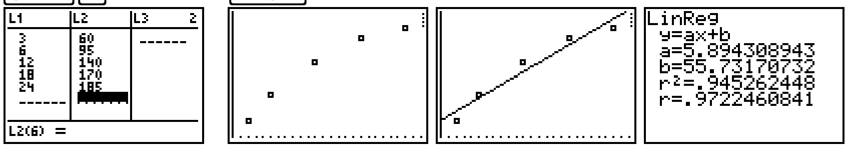
```
2-PropZTest
 p1<p2
 z=-3.194888227
 p=6.9948941E-4
 p̂1=.5132275132
 p̂2=.6201550388
↓p̂=.5749440716
```

```
2-PropZTest
 p1<p2
↑p̂1=.5132275132
 p̂2=.6201550388
 p̂=.5749440716
 n1=378
 n2=516
```

p = 6.99×10^{-4} = .000699 < .01 → Reject H_0.

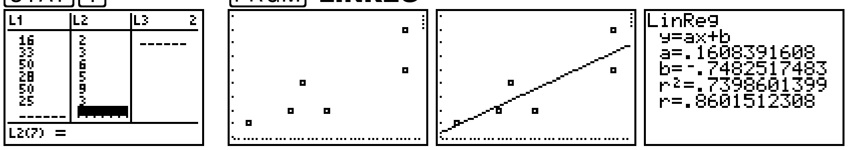
10.1 ---

596, 7 [STAT][1] [PRGM] **LINREG**

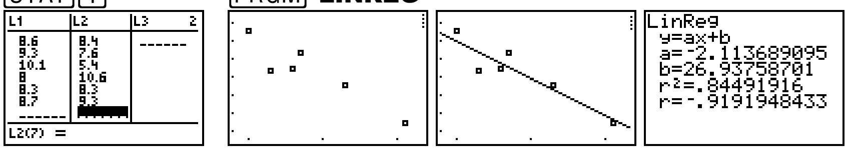


10.2 ---

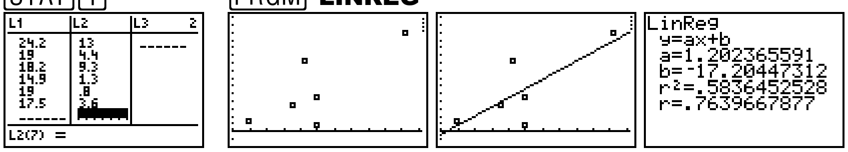
614, 1 [STAT][1] [PRGM] **LINREG**



616, 7 [STAT][1] [PRGM] **LINREG**

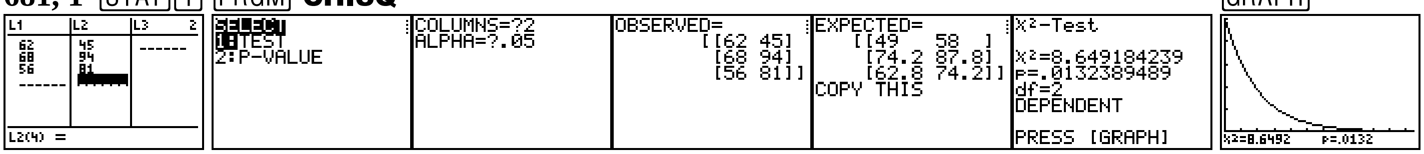


616, 9 [STAT][1] [PRGM] **LINREG**

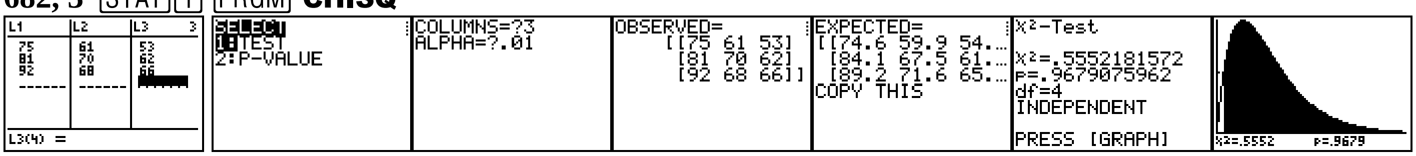


11.1 ---

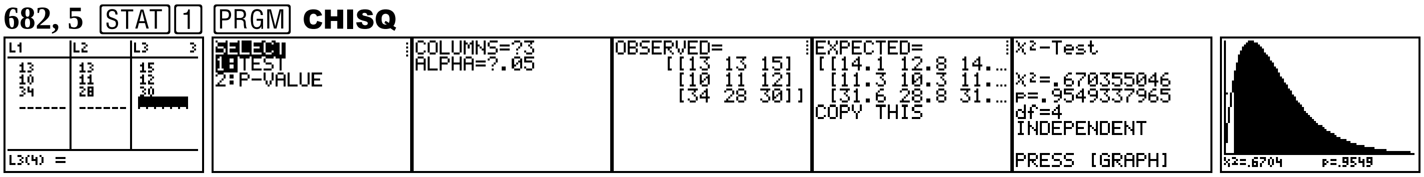
681, 1 [STAT][1] [PRGM] **CHISQ** [GRAPH]



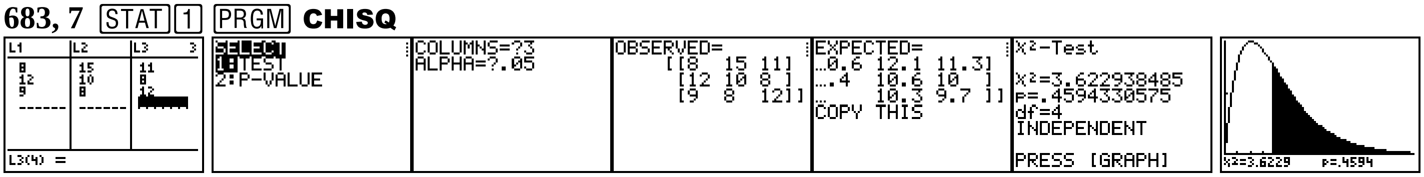
682, 3 [STAT][1] [PRGM] **CHISQ**



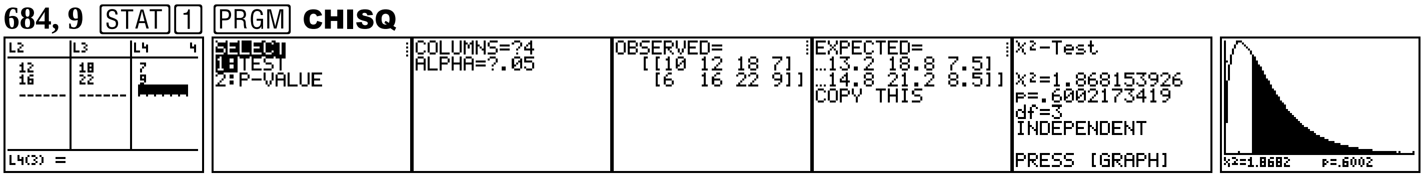
682, 5 [STAT][1] [PRGM] **CHISQ**



683, 7 [STAT][1] [PRGM] **CHISQ**



684, 9 [STAT][1] [PRGM] **CHISQ**



684, 9 STAT 1 PRGM **CHISQ**