Hypothesis Tests

Read Hypothesis Tests (How to Complete a Hypothesis Test)

Table of Contents for CALCULATORing
Table of Contents for MIDDLE GROUND, Statistics
TI83-84 STATISTICS and DISTRIBUTION Menus

Test	significance test for:	test statistic $(\text{test}) = \dfrac{(\text{observed"new"}) - (\text{expected"old"})}{(\text{standard error})}$
Z-Test	1 mean, normal, large sample size	$z_{test} = (\bar{x} - \mu)/(\sigma/\sqrt{(n)})$
T-Test	1 mean, small sample, sigma unknown	$t_{test} = (\bar{x} - \mu)/(s/\sqrt{(n)})$
2 Sample Z-Test	2 means	$z = \dfrac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1{}^2}{n_1} + \dfrac{\sigma_2{}^2}{n_2}}}$
2 Sample T-Test	2 INDEPENDENT means	$t = \dfrac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1{}^2}{n_1} + \dfrac{s_2{}^2}{n_2}}}$
2 Sample T-Test	2 DEPENDENT means	$t = \dfrac{\dfrac{\sum(x_1 - x_2)}{n} - \mu_{(x_1 - x_2)}}{\dfrac{\sqrt{\dfrac{n\sum(x_1 - x_2)^2 - (\sum(x_1 - x_2))^2}{n(n-1)}}}{\sqrt{n}}}$
ANOVA Test	3 or more means using variances	

		$F = \frac{s_{between}^2}{s_{within}^2}$
1 Proportion Z-Test	1 proportion	$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$
2 Proportion Z-Test	2 proportions	$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}(\frac{1}{n_1} + \frac{1}{n_2})}}$
χ^2 Test	discrete (catagories) Independence of Proportion Is the preference independent of the group?	$\chi^2 = \sum[\frac{(O-E)^2}{E}]$ CREATE an EXPECTED (row,column) distribution.
χ^2 Test	discrete (catagories) Homogeneity of Proportions Are preferences of subgroups the same?	$\chi^2 = \sum[\frac{(O-E)^2}{E}]$ CREATE an EXPECTED (row,column) distribution.
Linear Regression T-Test	1 correlation coefficient, between x and y	$t = r\sqrt{\frac{n-2}{1-r^2}}$
χ^2 GOF Test	Goodness of Fit, match "test" distribution to another distribution	$\chi^2 = \sum[\frac{(O-E)^2}{E}]$
χ^2 GOF Test	Normality, match "test" distribution to a normal distribution	$\chi^2 = \sum[\frac{(O-E)^2}{E}]$ CREATE an EXPECTED distribution.
χ^2 Test	1 Variance	$\chi^2_{test} = [(n-1)s^2]/\sigma^2$
Two Sample F Test	2 variances	$F = \frac{s_{\text{larger variance}}^2}{s_{\text{smaller variance}}^2}$

Sign Test	1 median or 1 pair, as in (After) - (Before) >0 not very sensitive a test	$z_{sign} = \frac{x+.5-n/2}{\sqrt{n}/2}$ $n = x^{+} + x^{-},\ n > 25$
Wilcoxon Mann-Whitney Rank Sum Test	2 distributions for symmetry, center, distribution of ranks with averaged ranks awarded to tied data points	$z_{Wilcoxon_{rank\ sum}} = \frac{R-\mu_R}{\sigma_R}$
Wilcoxon Signed-RankTest	median of 2 distributions before-after small sample size, not nec. normal	$z = \frac{W-0.5}{\sqrt{\frac{n(n+1)(2n+1)}{6}}}$ $W = \lvert \sum[\text{sgn}(x_2 - x_1) \cdot R] \rvert$

Z-Test Uses the Normal or Gaussian Probability Density Function or Distribution

This tests to see if evidence exists that the real mean is not the stated mean.
Needed:

$n \geq 30$
or, distribution is normal and standard deviation is known.

The null hypothesis is always:

H_0: μ = (stated mean of the population)

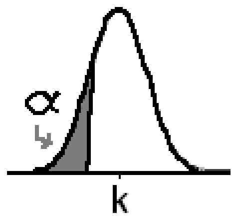
The alternate hypothesis is always one of these:

H_1: μ < (stated mean of the population)
H_1: μ = (stated mean of the population)
H_1: μ > (stated mean of the population)

Test statistic:

$z_{test} = (\bar{x} - \mu)/(\sigma/\sqrt{(n)})$

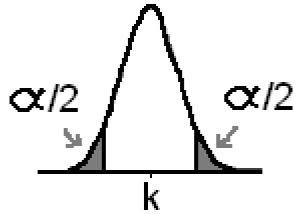
Levels of confidence:



$H_0: \mu = k$

$H_1: \mu < k$

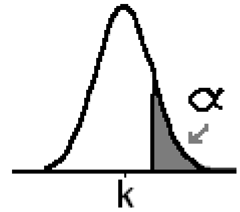
α	z critical
0.10	-1.28
0.05	-1.65
0.01	-2.33



$H_0: \mu = k$

$H_1: \mu \neq k$

α	z critical
0.10	± 1.65
0.05	± 1.96
0.01	± 2.58



$H_0: \mu = k$

$H_1: \mu > k$

α	z critical
0.10	1.28
0.05	1.65
0.01	2.33

To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then 7:Z-Interval...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then 1:Z-Test...
One page Summary, sample problem, and a spreadsheet.

T-Test

This tests to see if evidence exists that the real mean is not the stated mean.
Needed:
- approximately normal distribution,
- unknown population standard deviation.
- $n < 30$.
- the larger the sample size, the closed the distribution approaches normal.
- degrees of freedom, symbolized d.f., equals n-1.

The null hypothesis is always:
H_0: μ = (stated mean of the population)
The alternate hypothesis is always one of these:
H_1: μ < (stated mean of the population)
H_1: μ = (stated mean of the population)
H_1: μ > (stated mean of the population)
Test statistic:
$t_{test} = (\bar{x} - \mu)/(s/\sqrt{(n)})$
Levels of confidence:
found in table organized by 1-tail or 2-tail and d.f.
or by calculator
To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then 7:TInterval...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then 2:T-Test...
One page Summary, sample problem, and a spreadsheet.

1 Proportion Z-Test

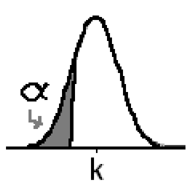
Used to test claims concerning percents, probabilities, proportions.
as in "70% of the population," "42% of the products"
Needed:
p, the population proportion, the stated proportion
$\hat{p}$ = x/n, the sample proportion
np≥5, nq≥5, where q= 1- p

$$\mu = n \bullet p \qquad \sigma = \sqrt{n \bullet p \bullet q}$$

The null hypothesis is always:
H_0: p = (stated proportion), may be symbolized as p_0
The alternate hypothesis is always one of these:
H_1: p ≠ (stated proportion)
H_1: p < (stated proportion)
H_1: p > (stated proportion)
Test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

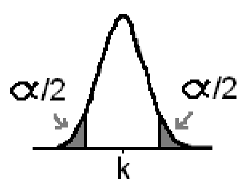
Levels of confidence: as above z-test



$H_0: \mu = k$

$H_1: \mu < k$

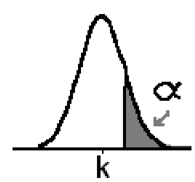
α	z critical
0.10	-1.28
0.05	-1.65
0.01	-2.33



$H_0: \mu = k$

$H_1: \mu \neq k$

α	z critical
0.10	± 1.65
0.05	± 1.96
0.01	± 2.58



$H_0: \mu = k$

$H_1: \mu > k$

α	z critical
0.10	1.28
0.05	1.65
0.01	2.33

or by calculator
To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then A:1-PropZInterval...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then 5:1-PropZTest...
One page Summary, sample problem, and a spreadsheet.

Linear Regression T-Test

This tests to see if evidence exists that there is a correlation between the independent variable, x, and the dependent variable, y. The populaltion correlation coefficient, rho,

symbol ρ, is the strength of the linear relationship between x and y.
The relationship ranges between
a strong negative relationship (as x increases, y decreases)
to a weak relationship, to no relationship (0),
to a weak positive relatioship (as x increases, y increases),
to a strong positive relationship.
$-1 \leq r \leq 1$

Needed:
r, the sample correlation coefficient is

$$r = \frac{n(\sum(xy)) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

d.f. = n-2
The null hypothesis is always:
H_0: $\rho = 0$
The alternate hypothesis is always one of these:
H_1: $\rho < 0$ H_1: $\rho \neq 0$ H_1: $\rho > 0$
Test statistic:

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

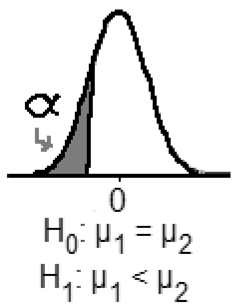
Levels of confidence:
found in table organized by 1-tail or 2-tail and d.f.
To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then G:LinRegTInt...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then F:LinRegTTest...
One page Summary, sample problem, and a spreadsheet.

2 Sample Z-Test

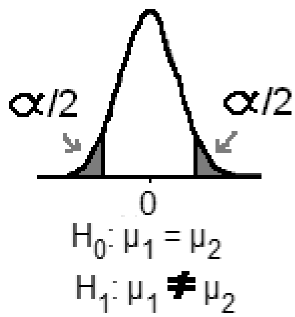
This tests to see if evidence exits that there is a difference between 2 means.
Needed:
two populations with known standard deviations.
independence between samples.
normal or approximately normal distributions if sample sizes are less than 30.
The null hypothesis is always:
H_0: $\mu_1 = \mu_2$ or H_0: $\mu_1 - \mu_2 = 0$
The alternate hypothesis is always one of these:
H_1: $\mu_1 \neq \mu_2$ or H_1: $\mu_1 - \mu_2 \neq 0$
H_1: $\mu_1 > \mu_2$ or H_1: $\mu_1 - \mu_2 > 0$
H_1: $\mu_1 < \mu_2$ or H_1: $\mu_1 - \mu_2 < 0$
Test statistic: [Because the null hypothesis is μ_1-μ_2=0, the numerator is simplified!]

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}}$$

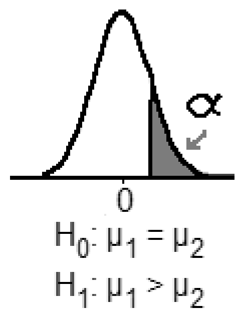
Levels of confidence:



α	z critical
0.10	-1.28
0.05	-1.65
0.01	-2.33



α	z critical
0.10	± 1.65
0.05	± 1.96
0.01	± 2.58



α	z critical
0.10	1.28
0.05	1.65
0.01	2.33

or by calculator

To Run a Calculator Confidence Interval to Test: (for 2-tail test)

[STAT], then [TESTS], then 9:2SampZInterval...

To Run a Calculator Hypothesis to Test:

[STAT], then [TESTS], then 3:2SampZ-Test...

One page Summary, sample problem, and a spreadsheet.

2 Sample T-Test, Independent Samples

This tests to see if evidence exits that there is a difference between 2 means.

Needed:

two populations, means and standard deviation need not be known.
sample standard deviations must be known.
independence between samples and are not matched or paired,
normal or approximately normal distributions if sample sizes are less than 30.

The null hypothesis is always:

H_0: $\mu_1 = \mu_2$ or H_0: $\mu_1 - \mu_2 = 0$

The alternate hypothesis is always one of these:

H_1: $\mu_1 \neq \mu_2$ or H_1: $\mu_1 - \mu_2 \neq 0$

H_1: $\mu_1 > \mu_2$ or H_1: $\mu_1 - \mu_2 > 0$

H_1: $\mu_1 < \mu_2$ or H_1: $\mu_1 - \mu_2 < 0$

Test statistic: [Because the null hypothesis is μ_1-μ_2=0, the numerator is simplified!]

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}}$$

Levels of confidence:

$$d.f. = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

d.f. smaller of n_1-1 and n_2-1
or by weighted degrees of freedom
as shown at right and in spreadsheet.

To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then 0:2-SampTInt...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then 4:2SampTTest...
One page Summary, sample problem, and a spreadsheet.

2 Sample T-Test, Dependent Samples, Paired, Before/After

$$\overline{D} = \frac{\sum(x_1 - x_2)}{n}$$

This tests hypothises about the mean of the difference of paired samples. It's great for before and after samples. Clearly the samples are not independent, but paired one-to-one to create differences which become the sample statistics.

This mean one is testing to see if there is a statistically significant difference before and after some treatment or event.

Needed:
$D = x_1 - x_2$ for every pair of n samples
they are n matched or paired samples, as in before and after samples
each population is normal or approximately normal
d.f. = n-1

The null hypothesis is always: H_0: $\mu_D = 0$ (, or some constant)
The alternate hypothesis is always one of these:
H_1: $\mu_D < 0$ (, or some constant)
H_1: $\mu_D \neq 0$ (, or some constant)
H_1: $\mu_D > 0$ (, or some constant)

Test statistic: [Because the null hypothesis is μ_1-μ_2=0 or some constant, the numerator may be simplified!]

$$t = \frac{\overline{D} - \mu_D}{\frac{\sqrt{\frac{n\sum D^2 - (\sum D)^2}{n(n-1)}}}{\sqrt{n}}} \qquad t = \frac{\frac{\sum(x_1 - x_2)}{n} - \mu_{(x_1 - x_2)}}{\frac{\sqrt{\frac{n\sum(x_1 - x_2)^2 - (\sum(x_1 - x_2))^2}{n(n-1)}}}{\sqrt{n}}}$$

Levels of confidence:

found in table organized by 1-tail or 2-tail and d.f.

To Run a Calculator Hypothesis to Test:

Create a list of the D distribution using $D = x_1 - x_2$ for every pair of n samples then [STAT], then [TESTS], then 2:T-Test... on this D distribution

One page Summary, sample problem, (calculator version), and a spreadsheet.

2 Proportion Z-Test

Used to test claims concerning 2 percents, probabilities, proportions, as in "70% of males," vs "42% of females."

Needed:

samples are independent

$n_1p_1 \geq 5$ $n_1q_1 \geq 5$ $n_2p_2 \geq 5$ $n_2q_2 \geq 5$

The null hypothesis is always:

H_0: $p_1 = p_2$

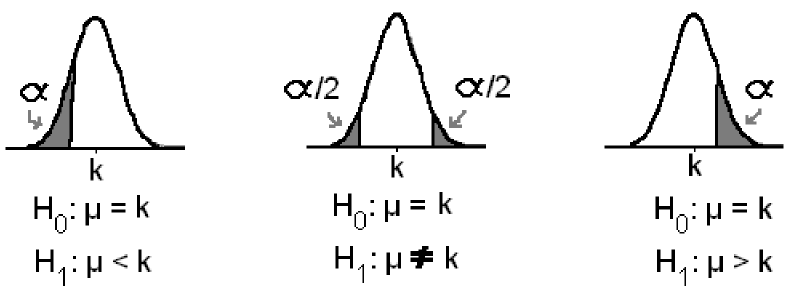
The alternate hypothesis is always one of these:

H_1: $p_1 > p_2$ H_1: $p_1 \neq p_2$ H_1: $p_1 < p_2$

Test statistic:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad z = \frac{\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) - (p_1 - p_2)}{\sqrt{\left(\frac{x_1 + x_2}{n_1 + n_2}\right)\left(1 - \frac{x_1 + x_2}{n_1 + n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Levels of confidence:



α	z critical
0.10	-1.28
0.05	-1.65
0.01	-2.33

α	z critical
0.10	± 1.65
0.05	± 1.96
0.01	± 2.58

α	z critical
0.10	1.28
0.05	1.65
0.01	2.33

To Run a Calculator Confidence Interval to Test: (for 2-tail test)

[STAT], then [TESTS], then B:2-PropInterval...

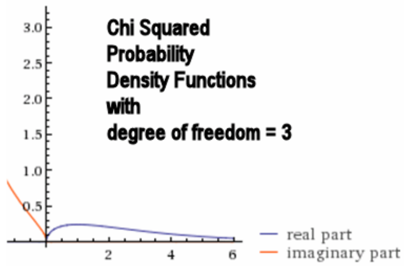
To Run a Calculator Hypothesis to Test:

[STAT], then [TESTS], then 6:2PropZ-Test...

One page Summary, sample problem, and a spreadsheet.

χ^2 Test for A Single Variance

This tests to see if evidence exists that



the real variance is not the stated variance.

It's a test for consistency, hoping for little, but enough, variation.

Needed:
- normal population
- random sample of n independent observations
- population standard deviation and sample standard deviation
- degrees of freedom, d.f. = n-1

The null hypothesis is always:

H_0: σ^2 = (stated population variance)

The alternate hypothesis is always one of these:

H_1: σ^2 < (stated population variance)

H_1: $\sigma^2 \neq$ (stated population variance)

H_1: σ^2 > (stated population variance)

Test statistic:

$\chi^2_{test} = [(n-1)s^2]/\sigma^2$

Levels of confidence:

found in table organized by 1-tail or 2-tail and d.f.

CAN'T Run a Calculator Hypothesis to Test w/o a program

CAN Use [CATALOG] and χ^2cdf(in a program

8:χ^2cdf(lower bound,upperbound,degrees of freedom)

One page Summary, sample problem, and a spreadsheet.

χ^2 Test for Independence of Proportions

This tests to see if evidence exists that the preferences of the groups is independent of the group making the choice.

Needed:
- Radom sample of full population is sorted by groups and their preferences, usually sorted into rows and columns and called a contingency table.
- each cell of table must be 5 or greater or groups/preferences must be pooled
- degrees of freedom, d.f. = ((rows)-1)((columns)-1)

The null hypothesis is always:

H_0: The preference is independent of the group.

The alternate hypothesis is always:

H_1: The preference is dependent on the group.

Test statistic: EXPECTED value table must be generated first using

the ratio:(row sum)(column sum)((grand sum) for each cell in table.

$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right]$$

Levels of confidence:
found in table organized by 1-tail or 2-tail and d.f.
To Run a Calculator Hypothesis Test:
[MATRIX] then edit OBSERVED and EXPECTED matrices,
then [STAT], then [TESTS], then C:χ-Test...
One page Summary, sample problem, and a spreadsheet.

χ^2 Test for Homogeneity of Proportions

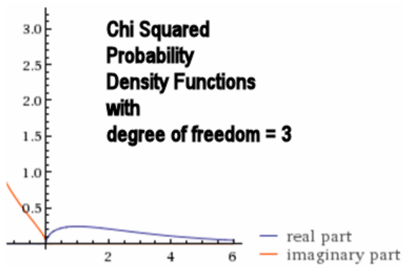
This tests to see if evidence exists that the one or more of the subgroup in a population has a diferent distribution of proportions than the other subgroups.

Needed:
each subgroup is randomly sampled,
preferences are sorted into rows and columns called a contingency table
each cell of table must be 5 or greater or groups/preferences must be pooled
degrees of freedom, d.f. = ((rows)-1)((columns)-1)
The null hypothesis is always:
H_0: $p_{subgroup1} = p_{subgroup2} = p_{subgroup3} = \dots = p_{last\ subgroup}$
The alternate hypothesis is always:
H_1: One or more of the ps is different.
Test statistic: EXPECTED value table must be generated first using the ratio:
(row sum)(column sum)((grand sum) for each cell in the contingency table.

$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right]$$

Levels of confidence:
found in table organized by 1-tail or 2-tail and d.f.
To Run a Calculator Hypothesis Test:
[MATRIX] then edit OBSERVED and EXPECTED matrices,
then [STAT], then [TESTS], then C:χ-Test...
One page Summary , sample problem, and a spreadsheet.

χ^2 Goodness of Fit Test



This tests to see if the observed result compares favorably to the theoretical (or prior) results.
It also tests if a distribution is normal.
The test looks at, all at once, how the frequency distributions of observed catagories compares to their expected frequencies.
It answers the question: "Does the observed result fit the pattern that was predicted?"
Recall that a completely random pattern would

have about equal frequencies in each catagory.

Needed:
- random sample of size n where n/(number of catagories) ≥ 5
- O = the observed frequency in each catagory
- E = the expected frequency in each catagory, WHICH IS...
  - based on a random distribution: E = n/(number of catagories)
  - based on a prior distribution: E = (prior catagory percent)(n)
- degrees of freedom, d.f. = (number of catagories) - 1

The null hypothesis is:
- H_0: The distribution is the same as the published distribution.
  (It's a reasonable fit.)
- H_0: The distribution is normally distributed.

The alternate hypothesis is:
- H_1: The distribution is not the same as the published distribution.
  (It's not a really good fit.)
- H_1: The distribution is not normally distributed.

Test statistic:

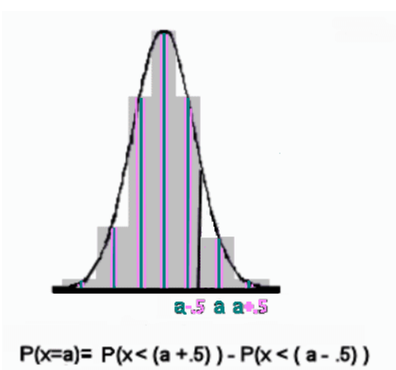
$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right]$$

Levels of confidence:
- found in table organized by 1-tail or 2-tail and d.f.

To Run a Calculator Hypothesis Test:
- Enter the OBSERVED *frequencies* in one list and EXPECTED in a matching list.
- [STAT], then [TESTS], then D:χ^2-Test...

One page Summary, sample problem, and a spreadsheet.

χ^2 Test for Normality



The Chi Square Goodness of Fit Test is here used to see if a distribution is normal.

Three steps are required.

1st: Compute the mean and standard deviation of the OBSERVED data (green).

2nd:Create an EXPECTED distribution using the normal cumulative distribution function and the OBSERVED mean and st.dev.(blue).

3rd: Run a GOF test using the OBSERVED and EXPECTED distribution frequencies (purple). The p in the test below is .619. The null hypothesis is confirmed. The OBSERVED bistribution is normal.

OBSERVED DISTRIBUTION

mean is 110.00
standard deviation is 34.66
variance 1201.50

	mark, x	O freq,f	fx	fx²
1	50	100	5000	250000
2	90	300	27000	2430000
3	130	300	39000	5070000
4	170	100	17000	2890000

mean is 110; Σfx is 88000
(Σfx)² is 7.7E+09
Σfx² is 1.1E+07

800 n
799 n-1 is

7.7E+08 n(Σfx²) - (Σfx)²
639200 n(n-1)

1201.5 [n(Σfx²) - (Σfx)²]/[n(n-1)], the variance
34.6627 √(variance), the standard deviation.

$$s^2 = \frac{n(\sum fx^2) - (\sum fx)^2}{n(n-1)}$$

EXPECTED DISTRIBUTION

OBSERVED statistics.
mean is 110.00
standard deviation is 34.66
variance 1201.50

	lower	upper	mark, x	Ex	Efrequ
1	30	70	50	0.1138	91.0030
2	70	110	90	0.3757	300.5964
3	110	150	130	0.3757	300.5964
4	150	190	170	0.1138	91.0030

	observed	expected	O - E	(O-E)^2	(O-E)^2/E
1	100	91.00	8.9970	80.9466	0.8895
2	300	300.60	-0.5964	0.3557	0.0012
3	300	300.60	-0.5964	0.3557	0.0012
4	100	91.00	8.9970	80.9466	0.8895

d.f is 3.00
chi squar 1.7814

COMPLETE A GOF TEST FOR NORMALITY OF DISTRIBUTION
H_0: Observed distribution is normal.
H_1: Observed distribution is not normal.

chisquaredGOFtest
p is 0.619

$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right]$$

One page Summary and a spreadsheet.

Two Sample F-Test

This tests to see if evidence exists that two variances differ.
Needed:

- 2 normal distributions
- independent samples
- sample standard deviations
- number in each sample
- d.f.n. (numerator) = $n_{numerator}-1$
- d.f.d. (denominator) = $n_{denominator}-1$

The null hypothesis is always:

H_0: $\sigma_{larger}^2 = \sigma_{smaller}^2$

The alternate hypothesis is always one of these:

H_1: $\sigma_{larger}^2 < \sigma_{smaller}^2$
H_1: $\sigma_{larger}^2 \neq \sigma_{smaller}^2$
H_1: $\sigma_{larger}^2 > \sigma_{smaller}^2$

Test statistic:

$$F = \frac{s_{\text{larger variance}}^2}{s_{\text{smaller variance}}^2}$$

Levels of confidence:
found in tables organized by d.f.n and d.f.d.

To Run a Calculator Confidence Interval to Test: (for 2-tail test)
[STAT], then [TESTS], then 7:TInterval...
To Run a Calculator Hypothesis to Test:
[STAT], then [TESTS], then 2:T-Test...
One page Summary, sample problem, and a spreadsheet.

ANOVA -- Analysis of Variance, One-Way, Single Factor

This tests to see if evidence exists that means of 3 or more groups differ when the variances are equal.
Needed:
population must be normal or near normal
independent samples
variances equal
$N = n_1 + n_2 + n_3 + ... + n_k$, k = the number of groups
d.f.numerator = k - 1
d.f.denominator = N - k
The grand mean, GM, is the mean of all data.
The null hypothesis is always:
H_0: $\mu_1 = \mu_2 = \mu_3 = ... = \mu_k$
The alternate hypothesis is always:
H_1: One or more of the means is not equal to the others
Test statistic:
The F test statistic,
the sum of the squares between groups is SS_B, $(SS_B)/(k-1) = MS_B = s_{between}^2$
the sum of the squares within groups is SS_W, $(SS_W)/(N-k) = MS_W = s_{within}^2$

$$F = \frac{s_{between}^2}{s_{within}^2} \quad s_{between}^2 = \frac{\sum[n_i(\bar{x}_i - \bar{x}_{GM})^2]}{k-1} \quad s_{within}^2 = \frac{\sum[(n_i - 1)s_i^2]}{\sum(n_i - 1)}$$

Levels of confidence:
found in table organized by alpha, d.f.numerator, d.f.denominator
To Run a Calculator Hypothesis to Test:
[STAT], then [EDIT] to enter lists, then [STAT], [TESTS], then H:ANOVA(L_1, L_2, L_3, ... ,L_{last})
Page one of Summary, page 2 of Summary , sample problem, and a spreadsheet.

Sign Test

This is really a binomial (+ or - response) test to see if evidence exists that a median is the hypothesized median (single sample) or to see if a difference exists in paired data. If n is less than 26, a formula is not used, just a count and compare procedure. If n is greater than 25, a z-test is used. In either case, the smaller of the number of data points above the median and number of data points below the mean is used to compute the test statistic.
Needed:
a sample of 100 to do the job a sample of 30 does w/a more efficient test
may be nonparametric (not defined mean, standard deviation, etc.)

random sample
use the raw data to find the "sign" of [(data)-(stated median)]
Compute each [(data)-(stated median)] or [(After)-(Before)] first to determine n = x+ + x- !!!

The null hypothesis always:

H_0: median = (stated median) or H_0: (After) - (Before) = 0

The alternate hypothesis is always one of these:

H_1: median < (stated median)	H_1: (After) - (Before) < 0
H_1: median = (stated median)	H_1: (After) - (Before) = 0
H_1: median > (stated median)	H_1: (After) - (Before) > 0

Test statistic:

When n < 26, n = sum of x+ and x-
smaller of:
x+, the # of data points greater than median (or 0)
and
x-, the # of data points less than the median (or 0)

binomial distribution used where p=.5,
n = the sum of x+ and x-.

or

When n > 25, n = sum of x+ and x-

$$z_{sign} = \frac{x + .5 - n/2}{\sqrt{n}/2}$$

$$n = x^+ + x^-, \quad n > 25$$

The z table is used with x in formula as the smaller x+ and x-.

One page Summary , sample problem, and a spreadsheet.

Wilcoxon Rank Sum Test

This tests to see if evidence exists that 2 distributions are the same by virtue of their symmetry, center, and the distribution of ranks (with averaged ranks awarded to tied data points). Do two populations have the same distribution, are their centers and spreads the same?

Needed:
2 not necessarily normal distributions
each sample size is 10 or greater
independent random samples

The null hypothesis is always:
H_0: The distributions are the same.
(The means and rank sums are "the same.")

Some alternate hypothesis are:
H_1: The distributions are not the same.
Are the means and rank sums close but not "the same?" (≠)
H_1: The first distributions is to the left of the other.
Is the first's rank sums much smaller, data points more to the left, than the second's? (<)
H_1: The first distributions is to the right of the same.
Is the first's rank sums much larger, data points more to the right, than the second's? (>)

Test statistic: The sample size determines the labels in the formulas.
The label n1 is given to the smaller sample size -- fewer numbers to add.
The data is then POOLED, with averaged ranks awarded to tied data points.
Individual r1 and r2 are computed by adding the ranks of the data points from the samples.
R is r1, the sum of the ranks of the sample with the smaller sample size.

$$z_{Wilcoxon_{rank\ sum}} = \frac{R - \mu_R}{\sigma_R} \qquad \mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} \qquad \sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

One page Summary , sample problem, and a spreadsheet.

Wilcoxon Signed-Rank Test

This tests to see if evidence exists that the medians of two distributions are the same, or before-after differences exist.

Needed:

- n samples paired as x1 - x2 = difference $\neq$ 0.
- the ranks of the pooled absolute values of the differences, tie ranks averaged.
- "signed-rank" products = (original sign)(pooled rank)
- the absolute value of the sum of the products of the (original sign)(pooled rank) of the negative differences.
- the absolute value of the sum of the products of the (original sign)(pooled rank) of the positive differences.
- the SMALLER of the absolute values of the sums of the + or - signed-rank products
- 2 not necessarily normal distributions
- independent random samples

The null hypothesis is always:

H_0: median1 = median2, or H_0: before - after = 0

Some alternate hypothesis are:

H_1: median1 $\neq$ median2, or H_1: before - after $\neq$ 0

H_1: median1 > median2, or H_1: before - after > 0

H_1: median1 < median2, or H_1: before - after < 0

Test statistic:

- if n < 30, use table for critical value and the SMALLER for the test statistic
- if n $\geq$ 30, use z distribution for critical value and formula with $\mu = n(n+1)/4$ for test statistic OR
- if n $\geq$ 30, use z distribution for critical value and the absolute value of the sum of all (sign)(rank) in formula with $\mu = .5$ for test statistic

$$z = \frac{W - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{6}}}$$

$$W = |\sum[\mathrm{sgn}(x_2 - x_1) \cdot R]|$$

$$z_{Wilcoxon_{signed-rank}} = \frac{w_s - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

One page Summary , sample problem, and a spreadsheet.

Spearman rank correlation coefficient

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Kruskal-Wallis test

$$H = \frac{12}{N(N+1)}\left(\frac{R_1^{\,2}}{n_1} + \frac{R_2^{\,2}}{n_2} + \frac{R_3^{\,2}}{n_3} + \ldots + \frac{R_k^{\,2}}{n_k}\right) - 3(N+1)$$