Number of Bee Stings

3	90	4	4	80	200	2	700
2	1000	200	3000	3	150	100	1
400	40	5	30	1000	100	200	200
1	100	150	1000	3	800	3	200
100	20	400	10	50	2	3	2000

sample mean $\overline{x} = 308.9$
sample standard deviation $s = 597.8$
sample median $= 95$
$Q_1 = 3.2$ $Q_2 = 200$

EXAMPLE [Testing Hypotheses about the Mean Number of Bee Stings]

A mean of 10 stings a week amounts to 520 a year. Do these data indicate that the population mean number of stings is different from 520 stings per year?

Test with $\alpha = 05$.

SOLUTION AND DISCUSSION

1. Hypotheses. We are seeking evidence in support of $\mu \neq 520$ so the hypotheses should be formulated as

$$H_0 : \mu = 520 \qquad H_1 : \mu \neq 520$$

2. Test Criterion The sample size is large, so the test statistic is

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{\overline{X} - 520}{S/\sqrt{n}}$$

The alternative hypothesis is two-sided so $R : \ Z \leq -c$ and $Z \geq c$.

3. Rejection region We are given $\alpha = .05$ so $\alpha/2 = .025$. From the normal table, $z_{.025} = 1.96$, so the rejection region is $R : \ Z \leq -1.96$ and $Z \geq 1.96$.

4. Calculation of test statistic. Since $\overline{x} = 308.9$ and $s = 597.8$, the observed value of the test statistic is

$$z = \frac{308.9 - 520}{597.8/\sqrt{40}} = -2.23$$

5. Conclusion Since the observed value of the test statistic lies within the rejection region. At level of significance .05, we reject $H_0 : \mu = 520$ in favor of the alternative hypothesis that the mean number of stings per year is different from 520.

Because the observed value -2.23 is well within the rejection region, we also calculate the $P-$value

$$P - \text{value} \quad = P[Z \leq -2.23] + P[Z \geq 2.23]$$

$$= 2(0.0129) = .026$$

MORE ON TESTS CONCERNING A POPULATION MEAN

An alternative hypothesis is called **two-sided** if it is of the form $H_1 : \mu \neq \mu_0$. The alternative $H_1 : \mu > \mu_0$ and the alternative $H_1 : \mu < \mu_0$ are each called **one-sided**.

The choice of the critical region is based on the form of the alternative hypothesis. For large samples, where the test statistic is $Z = \sqrt{n}(\bar{X} - \mu_0)/S$, the rejection(or critical) region is

$$Z < -z_{\alpha/2} \quad \text{or} \quad Z > z_{\alpha/2}$$

where $z_{\alpha/2}$ is the standard normal upper $100\alpha/2-th$ percentile of the standard normal distribution.

SMALL SAMPLE INFERENCES ABOUT A NORMAL POPULATION MEAN

When the population distribution is normal,

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

has a standard normal distribution for every sample size n.

Replacing σ by the sample standard deviation S leads to a random quantity

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

that is more variable than than a standard normal. It has student's t distribution with $n-1$ degrees of freedom. (Recall divisor for sample variance)

SMALL SAMPLE SIZE - TESTING NORMAL POPULATION MEAN

$$Test\ statistic\ \ T = \frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}$$

Alternative Hypothesis	*Reject Null Hypothesis if*
$\mu < \mu_0$	$T < -t_\alpha$
$\mu > \mu_0$	$T > t_\alpha$
$\mu \neq \mu_0$	$T < -t_{\alpha/2}$
	$or\ \ t > t_{\alpha/2}$

EXAMPLE The percent cockle in paper made for copy machines has a long run mean of 10.0 percent. Six measurements are made under new machine conditions that are thought to reduce the amount of cockle.

Test, with $\alpha = .05$, the mean has decreased.

The $n = 6$ observations

8.5 8.7 10.2 9.8 8.6 7.0

have $\bar{x} = 8.80$ and $s = 1.126$.

SOLUTION We are trying to establish that μ is less than the long run average for the old method $= 10.0 = \mu_0$.

Let μ be the mean(population) of cockle

on paper produced under the new conditions.

1. Hypothesis

null hypothesis	$H_0 : \mu \geq 10.0.$
alternative hypothesis	$H_1 : \mu < 10.0.$

2. level of significance $\alpha = .05$.

3. Test Statistic The *Test statistic* is

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 10.0}{S/\sqrt{6}}$$

4. Rejection Region We reject the null hypothesis if t is too small. There are $n-1=6-1=5$ degrees of freedom. From the $t-$table, $t_{.05}=2.015$ so we reject the null hypothesis if $t<-2.015$.

5. Calculation

$$t=\frac{\bar{X}-10.0}{S/\sqrt{n}}=\frac{8.80-10.0}{1.126/\sqrt{6}}=-2.610$$

6. Conclusion Since $t=-2.610$ is less than -2.015 we reject the null hypothesis at level of significance .05.

The $P-$value is $P(t<-2.610)=.024$ (by computer)

This solution assumes that cockle has a normal distribution.

LARGE SAMPLE INFERENCES CONCERNING A PROPORTION

We treat cases where the sample size is large. When the null hypothesis specifies a value for the population proportion having a characteristic,

$H_0 : p = p_0$

we use the specified value p_0 to calculate the standard deviation of $\widehat{p}$ under H_0. The test statistic is of the form

$$Z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

which is approximately standard normal when the sample size is large. Here $\widehat{p}$ is the sample proportion having the characteristic.

EXAMPLE Out of a random sample of 400 customers asked about the service they received on warranty repairs, 28 reported they were unhappy with some aspect of the service. We will assume that 400 is a very small fraction of the customers who received warranty repairs.

Can we say that the proportion p, in the whole customer base who would be unhappy with the service, is different from 0.1? Test with $\alpha = .05$.

SOLUTION AND DISCUSSION

1. Hypotheses. We are seeking evidence in support of $p \neq 0.1$ so the hypotheses should be formulated

null hypothesis	$H_0 : p = 0.1.$
alternative hypothesis	$H_1 : p \neq 0.1.$

2. level of significance $\alpha = .05$.

3. Test Statistic

For testing $H_0 : p = .1$, the test statistic is

$$Z = \frac{\widehat{p} - .1}{\sqrt{.1 \times .9/400}}$$

4. Rejection Region We reject the null hypothesis if Z is either too small or two large

$Z < -c$ or $Z > c$.

Since $\alpha/2 = .05/2 = .025$ and, from the normal table, $z_{.025} = 1.96$ we reject the null hypothesis if $Z < -1.96$ or if $Z > 1.96$.

5. Calculation

The sample proportion $\widehat{p} = \frac{28}{400} = .07$ and

$$Z = \frac{.07 - .1}{\sqrt{.1 \times .9/400}} = \frac{-.6}{.3} = -2.00$$

6. Conclusion Since $z = -2.00$ is less than -1.96 we reject the null hypothesis at level of significance .05. The population proportion is different from .1.