Tests of significance Key

Chapter 17

Hypotheses are always stated in terms of population parameters.
H$_0$ is a null hypothesis. H_0 is often a statement that no effect or no difference is present.
H$_a$ is an alternative hypothesis. H_a says that a parameter differs from its null value μ_0 in a specific direction (one-sided alternative) or in either direction (two-sided alternative).
Significance tests for the null hypothesis $H_0 : \mu = \mu_0$ concerning the unknown population (true) mean μ are based on the **one-sample z test statistic**:

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

The z test assumes an SRS of size n from the population with Normal distribution.
Notation $P(Z \leq z)$=$P(Z < z)$ means the probability that variable Z is less than value of the test statistic z.

H_0 :	$\mu = \mu_0$	$\mu = \mu_0$	$\mu = \mu_0$
H_a :	$\mu < \mu_0$	$\mu > \mu_0$	$\mu \neq \mu_0$
$P - value$	$P(Z \leq z)$	$P(Z \geq z)$	$2P(Z \geq \|z\|)$
Table A	area to the left of z	area to the right of z	2 (area to the right of $\|z\|$)

Small P-value indicate strong evidence against H_0. If the P-value is as small or smaller than a specified value α, the data are statistically significant at significance level α.

1. Dementia is the loss of the intellectual and social abilities severe enough to interfere with judgment, behavior, and daily functioning. Alzheimer's disease is the most common type of dementia. One article explored the experience and struggles of people diagnosed with dementia and their families. For a SRS of 21 people with early-onset dementia the mean age at diagnosis was 52.5 years. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean age at diagnosis of all people with early-onset dementia is less than 55 years old? Assume that the population has a normal distribution with standard deviation of 6.8 years.

$n = 21$ $\bar{x} = 52.5$
$\alpha = 0.05$ $\mu_0 = 55$ $\sigma = 6.8$

a) State your hypotheses.

$H_0: \mu = 55$

$H_a: \mu < 55$

b) Calculate the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{52.5 - 55}{\frac{6.8}{\sqrt{21}}} = -1.6847 \approx -1.68$$

c) Find the p-value.

$P(Z < -1.68) = 0.0465$

d) Write your conclusion in terms of the problem.

$0.0465 < \alpha = 0.05 \Rightarrow$ reject Ho, conclude Ha

at the 5% level of significance, there is evidence that the true mean age at diagnosis of all people with early-onset dementia is less than 55 years old.

2. Statistics can help decide the authorship of literary works. Sonnets by a certain Elizabethan poet are known to contain an average of 8.9 new words (words not used in the poets other works). The standard deviation of the number of new words is $\sigma = 2.5$. Now a manuscript with 6 new sonnets has come to light, and scholars are debating whether it is the poets work. The new sonnets contain an average of 10.2 words not used in the poets known works. We expect poems by another author to contain more new words. At the 1% significance level, is there evidence that the new sonnets are not by our poet?

a) State your hypotheses.

$\mu_0 = 8.9$ $\sigma = 2.5$ $n = 6$ $\bar{x} = 10.2$ (mean of 6 new sonnets, sample mean)

$\alpha = 0.01$

$H_0: \mu = 8.9$

$H_a: \mu > 8.9$

b) Calculate the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{10.2 - 8.9}{\frac{2.5}{\sqrt{6}}} = 1.2737 \approx 1.27$$

c) Find the p-value.

$$P(Z > 1.27) = 1 - P(Z < 1.27) = 1 - .8980 = 0.1020$$

value from TA for $z = 1.27$

d) Write your conclusion in terms of the problem.

$0.1020 > \alpha = 0.01 \Rightarrow$ fail to reject H_0, cannot conclude H_a

at the 1% level of significance, there is no evidence that the true mean number of new words in the new sonnets is greater than 8.9 words. No evidence that the new sonnets aren't by our poet.

3. According to the Bureau of Crime Statistics and Research of Australia, as reported on Lawlink, the mean length of imprisonment for motor-vehicle-theft offenders in Australia is 16.7 months. 100 randomly selected motor-vehicle-theft offenders in Sydney, Australia, had a mean length of imprisonment of 17.8 months. At the 10% significance level, do the data provide sufficient evidence to conclude that the mean length of imprisonment for motor-vehicle-theft offenders in Sydney differs from the national mean in Australia? Assume that the population standard deviation of the lengths of imprisonment for motor-vehicle-theft offenders in Sydney is 6 months.

$\mu_0 = 16.7$ $n = 100$ $\bar{x} = 17.8$

$\alpha = 0.10$ $\sigma = 6$

a) State your hypotheses.

$H_0: \mu = 16.7$

$H_a: \mu \neq 16.7$

b) State in words what the alternative hypothesis means.

The true mean length of imprisonment for motor-vehicle-theft offenders in sydney is not equal to 16.7 months.

c) Calculate the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{17.8 - 16.7}{\frac{6}{\sqrt{100}}} = 1.833 \approx 1.83$$

d) Find the p-value.

$$2P(Z > |1.83|) = 2(1 - P(Z < 1.83)) = 2(1 - .9664) = 0.0672$$

e) Write your conclusion in terms of the problem.

$0.0672 < \alpha = 0.10 \Rightarrow$ reject H_0, conclude H_a

at the 10% significance level, there is evidence that the true mean length of imprisgnment for motor-vehicle-theft offenders in Sydney differs from the national mean in Australia of 16.7 months.