# MODULE 18

# 1-SAMPLE Z-TEST

**Contents**

18.1 11-Steps of Hypothesis Testing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 135
18.2 1-Sample Z-Test Specifics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 136
18.3 1-Sample Z-Test in R . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137

A FOUNDATION FOR MAKING STATISTICAL INFERENCES was provided in Modules 13-17. Most of the material in Modules 15 and 17 is related to a 1-Sample Z-test, which is formalized in this module. Other specific hypothesis tests are in Modules 19-22.

## 18.1 11-Steps of Hypothesis Testing

Hypothesis testing is a rigorous and formal procedure for making inferences about a parameter from a statistic. The 11 steps listed below will help make sure that all aspects important to hypothesis testing are completed. These steps should be used for all hypothesis tests in this and ensuing modules.

1. State the rejection criterion ($\alpha$).
2. State the null and alternative hypotheses to be tested and define the parameter(s).
3. Identify (and explain why!) the hypothesis test to use (e.g., 1-Sample t, 2-sample t, etc.).
4. Collect the data (address study type and if randomization occurred).
5. Check all necessary assumptions (describe how you tested the validity).
6. Calculate the appropriate statistic(s).
7. Calculate the appropriate test statistic.
8. Calculate the p-value.
9. State your rejection decision about $H_0$.
10. Summarize your findings in terms of the problem.
11. **If $H_0$ was rejected**, compute and interpret an appropriate confidence region for the parameter.

The order of some of these steps is arbitrary. However Steps 1-3 **MUST** be completed before collecting data (Step 4). Further note that Step 11 will be completed only to provide a more definitive statement about the value of the parameter when the $H_0$ was rejected (i.e., if the parameter differs from the hypothesized value, then provide a range for which the actual parameter may exist).

## 18.2 1-Sample Z-Test Specifics

A 1-Sample Z-Test tests $H_0 : \mu = \mu_0$, where $\mu_0$ represents a specific value of $\mu$, when $\sigma$ is known. Other specifics of this test were discussed in previous modules and are summarized in Table 18.1.

Table 18.1. Characteristics of a 1-Sample Z-Test.

- **Hypothesis:** $H_0 : \mu = \mu_0$
- **Statistic:** $\bar{x}$
- **Test Statistic:** $Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$
- **Confidence Region:** $\bar{x} + Z^* \frac{\sigma}{\sqrt{n}}$
- **Assumptions:**
  1. $\sigma$ is known
  2. $n \geq 30$, $n \geq 15$ and the **population** is not strongly skewed, OR the **population** is normally distributed.
- **Use with:** Quantitative response, one group (or population), $\sigma$ known.

The only test that can possibly be confused with a 1-Sample Z-Test is a 1-Sample t-Test (Module 19), which tests the same null hypothesis but when $\sigma$ is unknown.

### 18.2.1 Example - Intra-class Travel

Below are the 11-steps (Section 18.1) for completing a full hypothesis test for the following situation:

> *A dean is interested in the average amount of time it takes for students to get from one class to another. In particular, she wants to determine if it takes more than 10 minutes, on average, to go between classes. In an effort to test this hypothesis, she collected a random sample of 100 intra-class travel times and found the mean to be 10.12 mins. Assume that it is known from previous studies that the distribution of intra-class times is symmetric with a standard deviation of 1.60 mins. Test the dean's hypothesis with $\alpha = 0.10$.*

1. $\alpha$=0.10.
2. $H_0 : \mu = 10$ mins vs. $H_A : \mu > 10$ mins, where $\mu$ is the mean time for ALL intra-class travel events.
3. A 1-Sample Z-Test is required because (i) a quantitative variable (intra-class travel time) was measured, (ii) individuals from one group (or population) is considered (students at the Dean's school), and (iii) $\sigma$ is thought to be known (=1.60 mins).
4. The data appear to be part of an observational study (the dean did not impart any conditions on the students) with a random selection of individuals.
5. (i) $n = 100 \geq 30$ and (ii) $\sigma$ is thought to be known (=1.60 mins).
6. $\bar{x}$=10.12.
7. $Z = \frac{10.12-10}{\frac{1.60}{\sqrt{100}}} = \frac{0.12}{0.16}$=0.75.
8. p-value=0.2266.
9. $H_0$ is not rejected because the p-value > $\alpha$=0.10.
10. It appears that the mean time for **all** intra-class travel events is not more than 10 minutes.
11. The confidence region is not computed when $H_0$ is not rejected.

**R Appendix:**

```
( distrib(10.12,mean=10,sd=1.60/sqrt(100),lower.tail=FALSE) )
```

## 18.3 1-Sample Z-Test in R

If raw data exist, the calculations for a 1-Sample Z-test can be efficiently computed with `z.test()`. This function requires the vector of quantitative data as the first argument, the hypothesized value for $\mu$ in `mu=`, and the known $\sigma$ in `sd=`. Additionally, the type of alternative hypothesis may be declared in `alt=`, where `alt="two.sided"` (the default), `alt="less"`, and `alt="greater"` correspond to the "not equals", "less than", and "greater than" hypotheses, respectively. Finally, the level of confidence may be given as a proportion (between 0 and 1) in `conf.level=` (which defaults to 0.95). The `z.test()` results may be assigned to an object and submitted to `plot()` to visualize the test statistic and p-value.

### 18.3.1 Body Temperature

Below are the 11-steps (Section 18.1) for completing a full hypothesis test for the following situation:

> *Machowiak et al. (1992) critically examined the belief that the mean body temperature is 98.6$^o$F by measuring the body temperatures in a sample of healthy humans. Use their data in BodyTemp.csv, with a supposedly known $\sigma = 0.63^o$F and $\alpha = 0.01$ to determine if the mean body temperature differs from 98.6$^o$F.*

1. $\alpha$=0.01.
2. $H_0 : \mu = 98.6^o$F vs. $H_A : \mu \neq 98.6^o$F, where $\mu$ is the mean body temperature for ALL healthy humans. [*Note that not equals was used because the researchers want to determine if the temperature is* ***different from*** $98.6^o$*F.*]
3. A 1-Sample Z-Test is required because (i) a quantitative variable (i.e., body temperature) was measured, (ii) individuals from one group (or population) is considered (i.e., healthy humans), and (iii) $\sigma$ is thought to be known (= 0.63$^o$F).
4. The data appear to be part of an observational study although this is not made clear in the background information. There is also no evidence that randomization was used.
5. (i) $n = 130 \geq 30$ and (ii) $\sigma$ is thought to be known (= 0.63$^o$F).
6. $\bar{x} = 98.25^o$F (Table 18.2).
7. $Z = -6.35$ (Table 18.2).
8. p-value$< 0.00005$ (Table 18.2).
9. Reject $H_0$ because p-value$< \alpha = 0.01$.
10. It appears that the mean body temperature of ALL healthy humans is less than 98.6$^o$F. [*Note that the test was for a difference but because* $\bar{x} < 98.6$ *this more specific conclusion can be made.*]
11. I am 99% confident that the mean body temperature ($\mu$) for ALL healthy humans is between 98.1 and 98.4$^o$F (Table 18.2).

Table 18.2. Results from 1-Sample Z-Test for mean body temperature.

```
z = -6.3482, n = 130.000, Std. Dev. = 0.630, Std. Dev. of the sample mean =
0.055, p-value = 2.178e-10
99 percent confidence interval:
 98.10690 98.39156
sample estimates:
mean of bt$temp
       98.24923
```

**R Appendix:**

```
bt <- read.csv("data/BodyTemp.csv")
( bt.z <- z.test(bt$temp,mu=98.6,sd=0.63,conf.level=0.99) )
```