Example 8
Z-test – One sample
Comparing sample and population means (two-tailed)
Setup

A researcher wants to know if elevation influences people's blood hemoglobin content. The researcher has collected blood samples from 64 adult males that live at high elevations and measured their blood hemoglobin content (data provided below g/100 ml). He wants to know if the sample mean differs significantly from the worldwide population mean of 15.80 g/100 ml with a population standard deviation of 2.00 g/100 ml.

Blood hemoglobin content (g/100ml)

16.55	19.68	12.29	16.22
14.13	15.56	15.28	16.42
15.93	14.85	17.40	16.54
14.45	18.16	18.03	19.43
17.83	15.50	18.22	15.16
14.11	14.60	16.66	16.38
16.87	10.97	17.83	14.77
16.98	17.91	14.97	17.28
17.77	17.98	16.43	15.64
20.10	18.30	19.84	16.83
15.78	19.00	14.58	20.02
15.83	16.07	17.62	16.19
22.34	15.95	19.24	14.65
19.82	15.90	17.09	16.90
19.42	18.01	11.76	15.40
18.04	21.89	16.49	18.15

Example 8

Z-test – One sample
Comparing sample and population means (two-tailed)
Solution

1. State your question: Does hemoglobin content of people that live at higher elevations differ from people in general?
    a. Is this a good scientific question? Definable, measurable, and controllable
    b. Identify your population: Hemoglobin content of people
    c. Identify your dependent variable: Hemoglobin content of blood
    d. Identify your independent variable: Elevation
2. State your hypothesis set
    a. Verbal hypothesis: People that live at high elevation will not have the same hemoglobin content as people in general.
    b. Statistical hypothesis (H_O, H_A):
        H_O: $x = \mu$, $x = 15.80$ g/100 ml Elevation does not influence the blood hemoglobin content of people.
        H_A: $x \neq \mu$, $x \neq 15.80$ g/100 ml People that live at high elevation will not have the same hemoglobin content as people in general.
    c. Is your hypothesis set exhaustive? Yes
    d. Is your hypothesis set exclusive? Yes
3. State your significance level: $\alpha = 0.05$
4. Select the appropriate test.
    a. Variable scale
        i. Dependent variable: Ratio
            - Converted or transformed? No
        ii. Independent variable: Nominal
            - Converted or transformed? No
    b. What information is given or available?
        i. Sample data
        ii. Parameters: Population mean and population standard deviation
    c. Number of samples: 1
    d. Are the data paired or unpaired? Not applicable
    e. What aspect of the variable do you want to compare? Central Tendency -- Means
    f. State the test to be used: Z-test – One sample
        i. Are assumptions met? Yes
            Random sample
            Independent samples
            Normally distributed data – Tested

Shapiro-Wilk normality test

data: HighElev$Hemoglobin
W = 0.9824, p-value = 0.494

Since p> 0.05, the data seems to be normally distributed.

5. Conduct your sampling
   We obtained hemoglobin content from blood samples from 64 adult males from high elevation sites. (Why just males? Why just adults?)
6. Graph the data

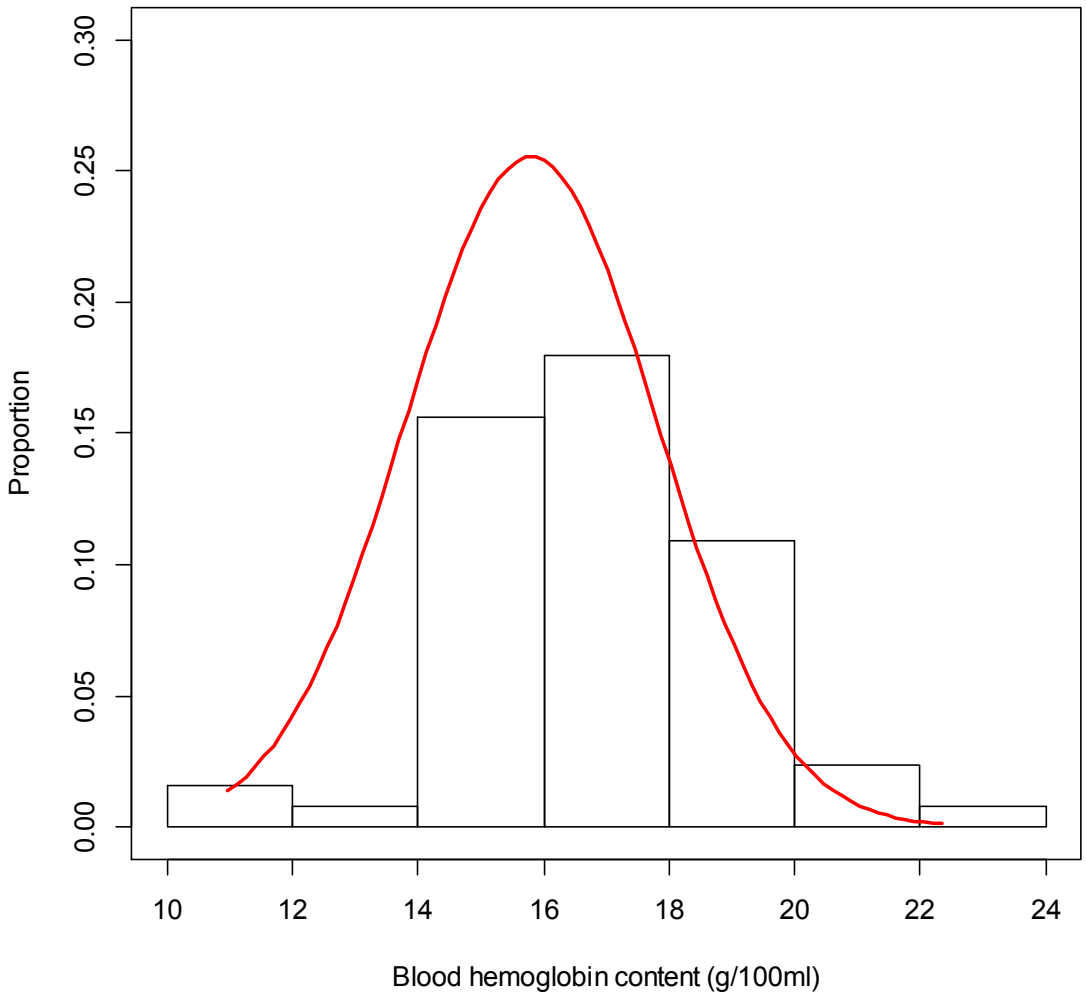

Figure 1. Blood hemoglobin content from 64 men that lived at high elevations. The red line represents the expected normal distribution (μ=15.80, σ=2.00).

7. Summarize the data

   Population parameters (Obtained from many years of research)

   μ = 15.80 g/100 ml

   σ = 2.00 g/100 ml

   Sample description

   n = 64

   x = 16.81 grams/100 ml

   Essentially we are asking does the 16.81 grams/100 ml differ significantly from 15.80 grams/100 ml?

8. Calculate your test statistic

   One Sample z-test

   data: SampleData
   z = 4.0494, n = 64.00, Std. Dev. = 2.00, Std. Dev. of the sample mean = 0.25, p-value = 5.135e-05
   alternative hypothesis: true mean is not equal to 15.8
   95 percent confidence interval:
   16.32235 17.30233
   sample estimates:
   mean of SampleData
   16.81234

9. Retain or reject your null hypothesis based on your test statistic

   The calculated p-value (5.135e-05) is less than the significance level (0.05), therefore we would reject our null hypothesis and retain our alternate hypothesis.

10. Interpret the results in biological terms.

    Blood hemoglobin content of people from high elevations differs from that of people in general ($Z=4.05$, $n=64$, $p<0.001$).