8.3: One-Sample Interval for the Mean

Suppose you want to estimate the mean height of Americans, or you want to estimate the mean salary of college graduates. A confidence interval for the mean would be the way to estimate these means.

Confidence Interval for One Population Mean (t-Interval)

1. State the random variable and the parameter in words.
   x = random variable
   μ = mean of random variable
2. State and check the assumptions for a hypothesis test
   a. A random sample of size n is taken.
   b. The population of the random variable is normally distributed, though the t-test is fairly robust to the assumption if the sample size is large. This means that if this assumption isn't met, but your sample size is quite large (over 30), then the results of the t-test are valid.
3. Find the sample statistic and confidence interval
   $\bar{x}-E<\mu<\bar{x}+E$
   where
   $E=t_c\frac{s}{\sqrt{n}}$
   $\bar{x}$ is the point estimator for μ
   t_c is the critical value where degrees of freedom: $df = n - 1$
   s is the sample standard deviation
   n is the sample size
4. Statistical Interpretation: In general this looks like, "there is a C% chance that the statement $\bar{x}-E<\mu<\bar{x}+E$ contains the true mean."
5. Real World Interpretation: This is where you state what interval contains the true mean.

The critical value is a value from the Student's t-distribution. Since a confidence interval is found by adding and subtracting a margin of error amount from the sample mean, and the interval has a probability of containing the true mean, then you can think of this as the statement $P(\bar{x}-E<\mu<\bar{x}+E)=C$. The critical values are found in table A.2 in the appendix.

How to check the assumptions of confidence interval:

In order for the confidence interval to be valid, the assumptions of the test must be true. Whenever you run a confidence interval, you must make sure the assumptions are true. You need to check them. Here is how you do this:

1. For the assumption that the sample is a random sample, describe how you took the sample. Make sure your sampling technique is random.
2. For the assumption that population is normal, remember the process of assessing normality from chapter 6.

Example $8.3.1$ confidence interval for the population mean using the formula

A random sample of 20 IQ scores of famous people was taken information from the website of IQ of Famous People ("IQ of famous," 2013) and then using a random number generator to pick 20 of them. The data are in Example $8.3.1$ (this is the same data set that was used in Example $8.3.2$). Find a 98% confidence interval for the IQ of a famous person.

Table $8.3.1$: IQ Scores of Famous People

158	180	150	137	109
225	122	138	145	180
118	118	126	140	165
150	170	105	154	118

1. State the random variable and the parameter in words.
2. State and check the assumptions for a confidence interval.

3. Find the sample statistic and confidence interval.
4. Statistical Interpretation
5. Real World Interpretation

Solution

1. x = IQ score of a famous person

μ = mean IQ score of a famous person

2.

a. A random sample of 20 IQ scores was taken. This was stated in the problem.
b. The population of IQ score is normally distributed. This was shown in Example $8.3.2$.

3. Sample Statistic:

$\overline{x} = 145.4$

$s \approx 29.27$

Now you need the degrees of freedom, *df* = *n* - 1 = 20 - 1 = 19 and the *C*, which is 98%. Now go to table A.2, go down the first column to 19 degrees of freedom. Then go over to the column headed with 98%. Thus $t_c = 2.539$. (See Example $8.3.2$.)

Degrees of Freedom (*df*)	80%	90%	95%	98%	99%
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
. . .	. . .	. . .	. . .	. . .	. . .
19	1.328	1.729	2.093	**2.539**	2.861

Table $8.3.2$: *Excerpt From Table A.2*

$$E = t_c \frac{s}{\sqrt{n}} = 2.539 \frac{29.27}{\sqrt{20}} \approx 16.6$$

$$\overline{x} - E < \mu < \overline{x} + E$$

$$145.4 - 16.6 < \mu < 145.4 + 16.6$$

$$128.8 < \mu < 162$$

4. There is a 98% chance that $128.8 < \mu < 162$ contains the mean IQ score of a famous person.

5. The mean IQ score of a famous person is between 128.8 and 162.

Example $8.3.2$ confidence interval for the population mean using technology

The data in Example $8.3.3$ are the life expectancies for men in European countries in 2011 ("WHO life expectancy," 2013). Find the 99% confident interval for the mean life expectancy of men in Europe.

Table $8.3.3$: Life Expectancies for Men in European Countries in 2011

7365	79	67	78	69	66	78	74
71	74	79	75	77	71	78	78
68	78	78	71	81	79	80	80
62	65	69	68	79	79	79	73
79	79	72	77	67	70	63	82
72	72	77	79	80	80	67	73
73	60	65	79	66			

1. State the random variable and the parameter in words.
2. State and check the assumptions for a confidence interval.
3. Find the sample statistic and confidence interval.
4. Statistical Interpretation
5. Real World Interpretation

Solution

1. x = life expectancy for a European man in 2011

μ = mean life expectancy for European men in 2011

2.

a. A random sample of 53 life expectancies of European men in 2011 was taken. The data is actually all of the life expectancies for every country that is considered part of Europe by the World Health Organization. However, the information is still sample information since it is only for one year that the data was collected. It may not be a random sample, but that is probably not an issue in this case.
b. The distribution of life expectancies of European men in 2011 is normally distributed. To see if this assumption has been met, look at the histogram, number of outliers, and the normal probability plot. (If you wish, you can look at the normal probability plot first. If it doesn't look linear, then you may want to look at the histogram and number of outliers at this point.)

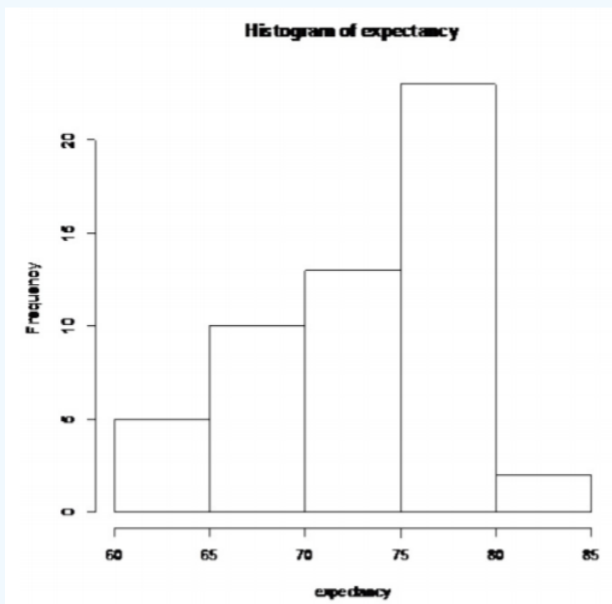

Figure 8.3.1: Histogram for Life Expectancies of European Men in 2011

Not normally distributed

Number of outliers:

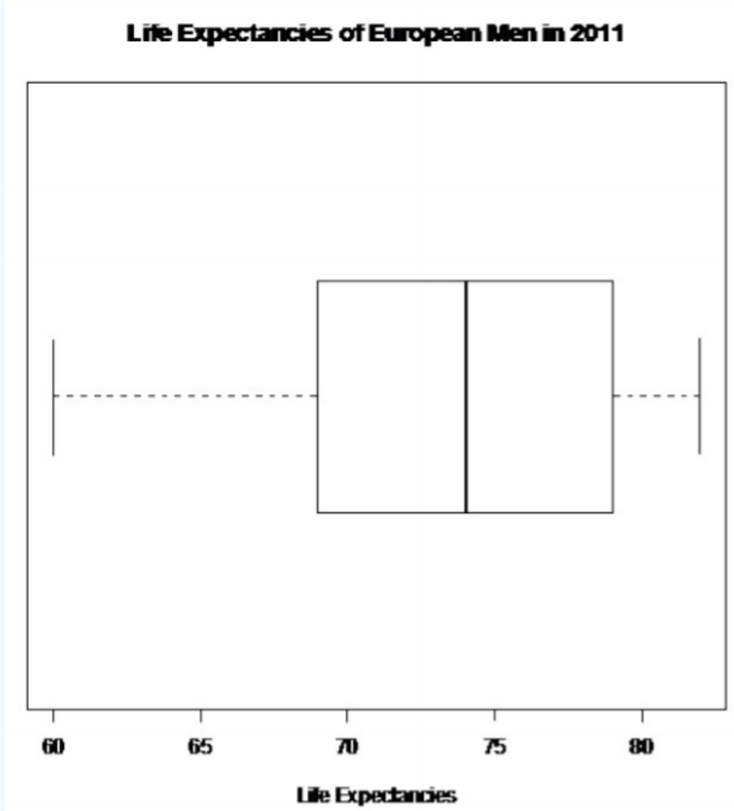

Figure 8.3.2: Modified Box Plot for Life Expectancies of European Men in 2011

IQR = 79 - 69 = 10

1.5 * IQR = 15

$Q1$ - 1.5 * IQR = 69 - 15 = 54

$Q3$ + 1.5 * IQR = 79 + 15 = 94

Outliers are numbers below 54 and above 94. There are no outliers for this data set.

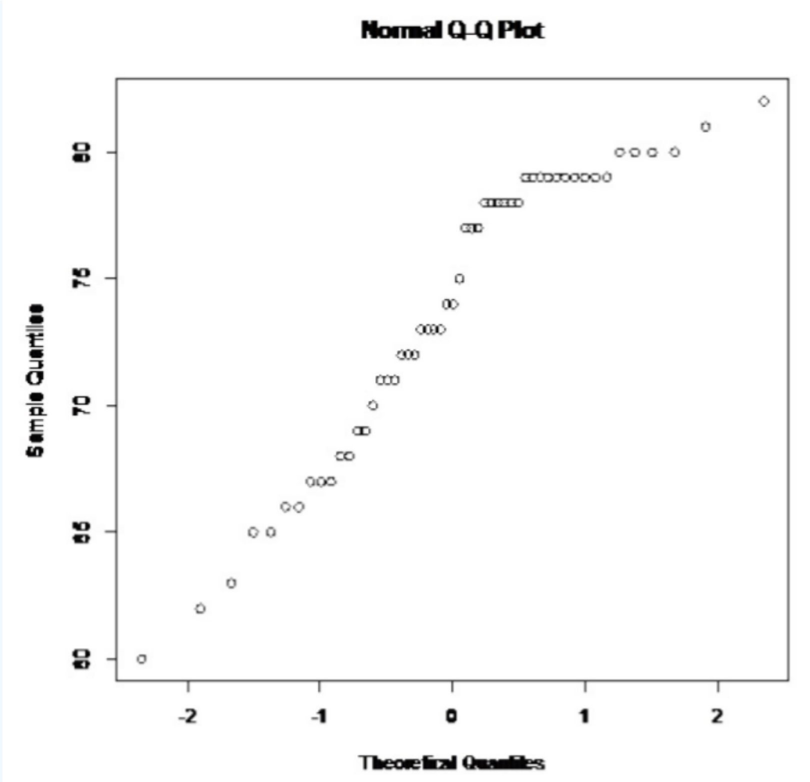

Figure 8.3.3: Normal Quantile Plot for Life Expectancies of European Men in 2011

Not linear

This population does not appear to be normally distributed. The t-test is robust for sample sizes larger than 30 so you can go ahead and calculate the interval.

3. Find the sample statistic and confidence interval

On the TI-83/84: Go into the STAT menu, and type the data into L1. Then go into STAT and over to TESTS. Choose TInterval.

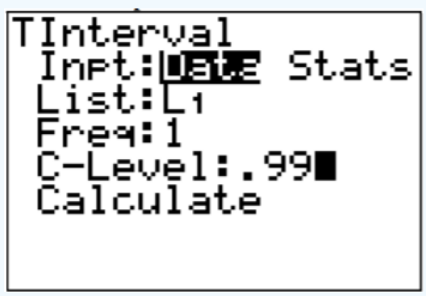

Figure 8.3.4: Setup for TInterval

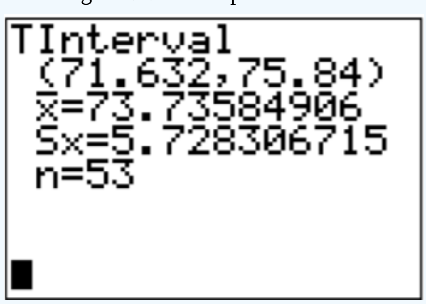

Figure 8.3.5: Results for TInterval

On R: t.test(variable, conf.level = C), where C is given in decimal form. So for this example it would be t.test(expectancy, conf.level = 0.99)

One Sample t-test

data: expectancy

t = 93.711, df = 52, p-value < 2.2e-16

alternative hypothesis: true mean is not equal to 0

99 percent confidence interval:

71.63204 75.83966

sample estimates:

mean of x

73.73585

71.6 years < μ 75.8 years

4. There is a 99% chance that 71.6 years < μ 75.8 years contains the mean life expectancy of European men.

5. The mean life expectancy of European men is between 71.6 and 75.8 years.

Homework

Exercise 8.3.1

In each problem show all steps of the confidence interval. If some of the assumptions are not met, note that the results of the interval may not be correct and then continue the process of the confidence interval.

1. The Kyoto Protocol was signed in 1997, and required countries to start reducing their carbon emissions. The protocol became enforceable in February 2005. Example 8.3.4 contains a random sample of CO2 emissions in 2010 ("CO2 emissions," 2013). Compute a 99% confidence interval to estimate the mean CO2 emission in 2010.

Table 8.3.4: CO2 Emissions (metric tons per capita) in 2010

1.36	1.42	5.93	5.36	0.06	9.11	7.32
7.93	6.72	0.78	1.80	0.20	2.27	0.28
5.86	3.46	1.46	0.14	2.62	0.79	7.48
0.86	7.84	2.87	2.45			

2. Many people feel that cereal is healthier alternative for children over glazed donuts. Example 8.3.5 contains the amount of sugar in a sample of cereal that is geared towards children ("Healthy breakfast story," 2013). Estimate the mean amount of sugar in children cereal using a 95% confidence level.

Table 8.3.5: Sugar Amounts (g) in Children's Cereal

10	14	12	9	13	13	13
11	12	15	9	10	11	3
6	12	15	12	12		

3. In Florida, bass fish were collected in 53 different lakes to measure the amount of mercury in the fish. The data for the average amount of mercury in each lake is in Example 8.3.6 ("Multi-disciplinary niser activity," 2013). Compute a 90% confidence interval for the mean amount of mercury in fish in Florida lakes.

Table 8.3.6: Average Mercury Levels (mg/kg) in Fish

1.23	1.33	0.04	0.44	1.20	0.27
0.48	0.19	0.83	0.81	0.81	0.5

0.49	1.16	0.05	0.15	0.19	0.77
1.08	0.98	0.63	0.56	0.41	0.73
0.34	0.59	0.34	0.84	0.50	0.34
0.28	0.34	0.87	0.56	0.17	0.18
0.19	0.04	0.49	1.10	0.16	0.10
0.48	0.21	0.86	0.52	0.65	0.27
0.94	0.40	0.43	0.25	0.27	

4. In 1882, Albert Michelson collected measurements on the speed of light ("Student t-distribution," 2013). His measurements are given in Example $8.3.7$. Find the speed of light value that Michelson estimated from his data using a 95% confidence interval.

Table $8.3.7$: Speed of Light Measurements in (km/sec)

299883	299816	299778	299796	299682
299711	299611	299599	300051	299781
299578	299796	299774	299820	299772
299696	299573	299748	299748	299797
299851	299809	299723		

5. Example $8.3.8$ contains pulse rates after running for 1 minute, collected from females who drink alcohol ("Pulse rates before," 2013). The mean pulse rate after running for 1 minute of females who do not drink is 97 beats per minute. Do the data show that the mean pulse rate of females who do drink alcohol is higher than the mean pulse rate of females who do not drink? Test at the 5% level.

Table $8.3.8$: Pulse Rates of Woman Who Use Alcohol

176	150	150	115	129	160
120	125	89	132	120	120
68	87	88	72	77	84
92	80	60	67	59	64
88	74	68			

6. The economic dynamism, which is the index of productive growth in dollars for countries that are designated by the World Bank as middle-income are in Example $8.3.9$ ("SOCR data 2008," 2013). Countries that are considered high-income have a mean economic dynamism of 60.29. Do the data show that the mean economic dynamism of middle-income countries is less than the mean for high-income countries? Test at the 5% level.

Table $8.3.9$: Economic Dynamism ($) of Middle Income Countries

25.8057	37.4511	51.915	43.6952	47.8506	43.7178	58.0767
41.1648	38.0793	37.7251	39.6553	42.0265	48.6159	43.8555
49.1361	61.9281	41.9543	44.9346	46.0521	48.3652	43.6252
50.9866	59.1724	39.6282	33.6074	21.6643		

7. In 1999, the average percentage of women who received prenatal care per country is 80.1%. Example 8.3.10 contains the percentage of woman receiving prenatal care in 2009 for a sample of countries ("Pregnant woman receiving," 2013). Do the data show that the average percentage of women receiving prenatal care in 2009 is higher than in 1999? Test at the 5% level.

Table 8.3.10: Percentage of Woman Receiving Prenatal Care

70.08	72.73	74.52	75.79	76.28	76.28
76.65	80.34	80.60	81.90	86.30	87.70
87.76	88.40	90.70	91.50	91.80	92.10
92.20	92.41	92.47	93.00	93.20	93.40
93.63	93.69	93.80	94.30	94.51	95.00
95.80	95.80	96.23	96.24	97.30	97.90
97.95	98.20	99.00	99.00	99.10	99.10
100.00	100.00	100.00	100.00	100.00	

8. Maintaining your balance may get harder as you grow older. A study was conducted to see how steady the elderly is on their feet. They had the subjects stand on a force platform and have them react to a noise. The force platform then measured how much they swayed forward and backward, and the data is in Example 8.3.11 ("Maintaining balance while," 2013). Do the data show that the elderly sway more than the mean forward sway of younger people, which is 18.125 mm? Test at the 5% level.

Table 8.3.11: Forward/Backward Sway (in mm) of Elderly Subjects

19	30	20	19	29	25	21	24	50

Answer

For all confidence intervals, just the interval using technology is given. See solution for the entire answer.

1. $1.7944 < \mu < 5.1152$ metric tons per capita

3. $0.44872 < \mu < 0.60562$ mg/kg

5. $87.2423 < \mu < 113.795$ beats/min

7. $88.8747\% < \mu < 93.0253\%$

Data Sources:

Australian Human Rights Commission, (1996). *Indigenous deaths in custody 1989 - 1996.* Retrieved from website: www.humanrights.gov.au/public...deaths-custody

CDC features - new data on autism spectrum disorders. (2013, November 26). Retrieved from www.cdc.gov/features/countingautism/

Center for Disease Control and Prevention, Prevalence of Autism Spectrum Disorders - Autism and Developmental Disabilities Monitoring Network. (2008). *Autism and developmental disabilities monitoring network-2012.* Retrieved from website: www.cdc.gov/ncbddd/autism/doc...nityReport.pdf

CO2 emissions. (2013, November 19). Retrieved from http://data.worldbank.org/indicator/EN.ATM.CO2E.PC

Federal Trade Commission, (2008). *Consumer fraud and identity theft complaint data: January-december 2007.* Retrieved from website: www.ftc.gov/opa/2008/02/fraud.pdf

Gallup news service. (2013, November 7-10). Retrieved from www.gallup.com/file/poll/1658...acy_131115.pdf

Healthy breakfast story. (2013, November 16). Retrieved from lib.stat.cmu.edu/DASL/Stories...Breakfast.html

Maintaining balance while concentrating. (2013, September 25). Retrieved from http://www.statsci.org/data/general/balaconc.html

Morgan Gallup poll on unemployment. (2013, September 26). Retrieved from http://www.statsci.org/data/oz/gallup.html

Multi-disciplinary niser activity - mercury in bass. (2013, November 16). Retrieved from http://gozips.uakron.edu/~nmimoto/pa.../MercuryInBass - description.txt

Pregnant woman receiving prenatal care. (2013, October 14). Retrieved from http://data.worldbank.org/indicator/SH.STA.ANVC.ZS

Pulse rates before and after exercise. (2013, September 25). Retrieved from http://www.statsci.org/data/oz/ms212.html

SOCR data 2008 world countries rankings. (2013, November 16). Retrieved from wiki.stat.ucla.edu/socr/index...ountriesRankin gs

Student t-distribution. (2013, November 25). Retrieved from lib.stat.cmu.edu/DASL/Stories/student.html

WHO life expectancy. (2013, September 19). Retrieved from www.who.int/gho/mortality_bur...n_trends/en/in dex.html

---