probability and statistics morgan bergen

10 one and two sample tests of hypotheses may 1 2023

statistical hypotheses general concepts	tests on a single mean variance unknown
single sample tests concerning a single mean	inferences on paired samples
tests on a single mean variance known	one sample test on a single proportion
relationship to confidence interval estimation	two samples test on two proportions

10.22 In the American Heart Association journal Hypertension, researchers report that individuals who practice Transcendental Meditation (TM) lower their blood pressure significantly.

If a random sample of 225 male TM practitioners meditate for 8.5 hours per week with a standard deviation of 2.25 hours, does that suggest that, on average, men who use TM meditate more than 8 hours per week?

Quote a P-value in your conclusion.

ASSUME THE LEVEL OF SIGNIFICANCE IS $\alpha = 0.05$

μ WILL REPRESENT THE TRUE POPULATION MEAN OF MEDITATION TIME FOR MALES

NUMBER OF MALES $N = 225$

MEAN $\bar{x} = 8.5$

STANDARD DEVIATION $S = 2.25$

NULL HYPOTHESIS $H_0 : \mu = 8$

ALTERNATIVE HYPOTHESIS $H_1 : \mu > 8$

CORRESPONDING Z VALUE

$$z = \frac{\bar{x} - \mu_0}{S/\sqrt{N}} = \frac{8.5 - 8}{2.25/\sqrt{225}} = \frac{0.5}{0.15} \simeq 3.33$$

NORMAL TABLE WE CAN FIND THE P-VALUE AS FOLLOWS

$$P = P(Z > z) = 1 - P(Z \leq 3.33) = 1 - 0.9996 = 0.0004$$

BECAUSE $0.0004 < 0.05$ == PVALUE < SIGNIFICANCE LEVEL WE CAN REJECT THE NULL HYPOTHESIS & CONCLUDE THAT THE MEAN MEDITATION TIME FOR MEN IS MORE THAN 8 HOURS PER WEEK.

probability and statistics morgan bergen

10 one and two sample tests of hypotheses may 1 2023

statistical hypotheses general concepts	tests on a single mean variance unknown
single sample tests concerning a single mean	inferences on paired samples
tests on a single mean variance known	one sample test on a single proportion
relationship to confidence interval estimation	two samples test on two proportions

10.45 a taxi company manager is trying to decide whether the use of radial tires instead of regular belted tires improves fuel economy. Twelve cars were equipped with radial tires and driven over a prescribed test course. without changing drivers, the same cars were then equipped with regular belted tires and driven once again over the test course. the gasoline consumption, in kilometers per liter, was recorded as follows in the following table. can we conclude that cars equipped with radial tires give better fuel economy than those equipped with belted tires? assume the populations to be normally distributed. Use a P-value in your conclusion.

kilometers per liter

car	**radial tire**	**belted tires**
1	4.2	4.1
2	4.7	4.9
3	6.6	6.2
4	7.0	6.9
5	6.7	6.8
6	4.5	4.4
7	5.7	5.7
8	6.0	5.8
9	7.4	6.9
10	4.9	4.7
11	6.1	6.0
12	5.2	4.9

$N = 12$

$\alpha = 0.05$

μ_1 POPULATION MEAN FOR RADIAL TIRES

μ_2 POPULATION MEAN FOR BELTED TIRES

$H_0 : \mu_1 - \mu_2 = 0$ NULL HYPOTHESIS

$H_1 : \mu_1 - \mu_2 > 0$ ALTERNATIVE HYPOTHESIS

$d_1 = 4.2 - 4.1 = 0.1$

$d_2 = 4.7 - 4.9 = -0.2$

$d_3 = 6.6 - 6.2 = 0.4$

$d_4 = 7.0 - 6.9 = 0.1$

$d_5 = 6.7 - 6.8 = -0.1$

$d_6 = 4.5 - 4.4 = 0.1$

$d_7 = 5.7 - 5.7 = 0$

$d_8 = 6.0 - 5.8 = 0.2$

$d_9 = 7.4 - 6.9 = 0.5$

$d_{10} = 4.9 - 4.7 = 0.2$

$d_{11} = 6.1 - 6.0 = 0.1$

$d_{12} = 5.2 - 4.9 = 0.3$

$$\bar{d} = \frac{0.1 + 0.2 + 0.4 + \cdots + 0.3}{12} = 0.14$$

$$s^2_d = \frac{1}{N-1} \sum_{i=1}^{N=12} (d_i - \bar{d})^2 = \frac{1}{12} \left((0.1 - 0.14)^2 + \cdots + (0.3 - 0.14)^2 \right)$$

$s^2_d = 0.04$

$s_d = 0.2$

$$t = \frac{\bar{d} - d_0}{s_d / \sqrt{N}} = \frac{0.14 - 0}{0.2 / \sqrt{12}} = 2.48$$

DEGREES OF FREEDOM $12 - 1 = 11 = df$

PVALUE $= 0.0153 < 0.05 =$ LEVEL OF SIG

$\therefore$ CARS WITH RADIAL TIRES HAVE A BETTER FUEL ECONOMY

10.23 add p-value approach

10.23 test the hypothesis that the average content of containers of a particular lubricant is 10 liters if the contents of a random sample of 10 containers are 10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3, and 9.8 liters. use a 0.01 level of significance and assume that the distribution of contents is normal.

SAMPLE SIZE $N = 10$

NULL HYPOTHESIS $H_0 : \mu = 10$

ALTERNATIVE HYPOTHESIS $H_1 : \mu \neq 10$

SAMPLE MEAN

$$\bar{x} = 1/N \sum_{i=1}^{N} x_i = 1/10 \cdot 100.6 = 10.06$$

SAMPLE STANDARD DEVIATION

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2} = \sqrt{\frac{1}{9} \cdot 0.5440} = 0.246$$

TEST STATISTIC

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{N}} \sim t_{(N-1)}$$

CRITICAL REGION

$$t > t_{\alpha/2,(N-1)} \quad \text{OR} \quad t < -t_{\alpha/2,(N-1)}$$

$$t = \frac{10.06 - 10}{0.246/\sqrt{10}} = \frac{0.06}{0.078} = 0.77$$

t VALUE WITH $10 - 1 = 9$ DEGREES OF FREEDOM $\alpha = 0.01$ LEVEL OF SIGNIFICANCE

CRITICAL VALUE,

$$t_{0.005,9} = 3.250$$

$t < 3.250$, WE DONT REJECT THE NULL HYPOTHESIS

THEREFORE THE CONCLUSION IS THAT THE AVERAGE CONTENT OF THE LUBRICANT IS 10 LITERS

10.24 the average height of females in the freshman class of a certain college has historically been 162.5 centimeters with a standard deviation of 6.9 centimeters. is there reason to believe that there has been a change in the average height if a random sample of 50 females in the present freshman class has an average height of 165.2 centimeters?
use a P-value in your conclusion. assume the standard deviation remains the same.

LEVEL OF SIGNIFICANCE $\alpha = 0.05$

$\alpha/2 = 0.05/2 = 0.025$

NULL HYPOTHESIS $H_0 : \mu = 162.5$

ALTERNATIVE HYPOTHESIS $H_1 : \mu \neq 162.5$

THE CRITICAL REGION

$Z(\text{OBSERVED VALUE}) < -Z_{\alpha/2}$ OR $Z(\text{OBSERVED VALUE}) > Z_{\alpha/2}$

$Z_{\alpha/2}$ CAN BE FOUND FROM THE Z TABLE NORMAL DISTRIBUTION

TO FIND THE CORRESPONDING Z VALUE HAS THE FOLLOWING FORMULA

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{N}} = \frac{165.2 - 162.5}{6.9/\sqrt{50}} \approx 2.77$$

$\bar{x} = 165.2$ SAMPLE MEAN — AVG HEIGHT FROM SAMPLE

$\mu_0 = 162.5$ MEAN — AVG HEIGHT OF FEMALES

$\sigma = 6.9$ STANDARD DEVIATION

$N = 50$ TOTAL NUMBER OF RANDOM SAMPLES

$Z(\text{OBSERVED VALUE}) = 2.77$

$Z_{\alpha/2} = Z_{0.025} = 1.96$

CRITICAL REGION IS THUS,

$Z(\text{OBSERVED VALUE}) < -Z_{0.025} = -1.96$ OR $Z(\text{OBSERVED VALUE}) > Z_{0.025} = 1.96$

THEREFORE WE REJECT H_0 & CONCLUDE THAT AVERAGE HEIGHT IS NOT 162.5

P-VALUE is $P(|Z| > 2.77) = 2P(Z < -2.77)$

$= 2 \cdot 0.0028$

$= 0.0056$

PVALUE < LEVEL OF SIGNIFICANCE

$0.0056 < 0.05$

THEREFORE WE REJECT H_0 & CONCLUDE THAT
THAT THERE HAS NOT BEEN A CHANGE IN THE AVG HEIGHT

probability and statistics morgan bergen

10 one and two sample tests of hypotheses may 1 2023

statistical hypotheses general concepts	tests on a single mean variance unknown
single sample tests concerning a single mean	inferences on paired samples
tests on a single mean variance known	one sample test on a single proportion
relationship to confidence interval estimation	two samples test on two proportions

10.26 according to a dietary study, high sodium in- take may be related to ulcers, stomach cancer, and migraine headaches. the human requirement for salt is only 220 milligrams per day, which is surpassed in most single servings of ready-to-eat cereals. if a random sample of 20 similar servings of a certain cereal has a mean sodium content of 244 milligrams and a standard deviation of 24.5 milligrams, does this suggest at the 0.05 level of significance that the average sodium content for a single serving of such cereal is greater than 220 milligrams? assume the distribution of sodium contents to be normal.

$\mu_0 = 220$ HUMAN REQUIREMENT 220 MILIGRAMS PER DAY

$N = 20$ RANDOM SAMPLE OF 20 SERVINGS

$\bar{X} = 244$ MEAN SODIUM CONTENT IS 244 MILIGRAMS PER DAY

$S = 24.5$ STANDARD DEVIATION OF 24.5 MILIGRAMS

$\alpha = 0.05$ LEVEL OF SIGNIFICANCE OF 0.05

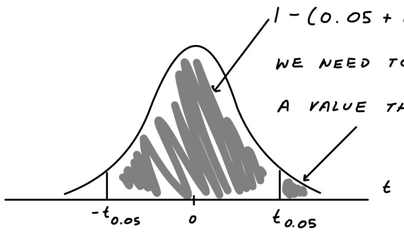
$H_1 : \mu > 220$ AVG IS GREATER THAN 220 MILIGRAMS ?

$H_0 : \mu = 220$ NULL HYPOTHESES

$H_1 : \mu > 220$ ALTERNATIVE HYPOTHESIS

TEST STATISTIC

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{N}} = \frac{244 - 220}{24.5/\sqrt{20}} \approx 4.38$$



$1 - (0.05 + 0.05) = 1 - 0.1 = 0.9$

WE NEED TO FIND THIS T VALUE

A VALUE THAT LEAVES AN AREA OF 0.05 TO THE RIGHT

FROM THE t-TABLE WITH 19 DEGREES OF FREEDOM

$t_{0.05} = 1.729$

BECAUSE, $t = 4.3808 > t_{0.05} = 1.729$

WE REJECT THE NULL HYPOTHESIS OF $H_0 : \mu = 220$

WE ACCEPT THE ALTERNATIVE HYPOTHESIS OF $H_1 : \mu > 220$

THUS THE AVERAGE CONTENT FOR CEREAL IS GREATER THAN 220 MILIGRAMS

10.42 five samples of a ferrous-type substance were used to determine if there is a difference between a laboratory chemical analysis and an X-ray fluorescence analysis of the iron content.

each sample was split into two subsamples and the two types of analysis were applied. following are the coded data showing the iron content analysis:

assuming that the populations are normal, test at the 0.05 level of significance whether the two methods of analysis give, on the average, the same result.

	sample				
analysis	1	2	3	4	5
x-ray	2.0	2.0	2.3	2.1	2.4
chemical	2.2	1.9	2.5	2.3	2.4

GIVEN INFORMATION,

$N = 5$ 5 TOTAL RANDOM SAMPLES
$\alpha = 0.05$ LEVEL OF SIGNIFICANCE IS 0.05
μ_1 DENOTES THE POPULATION AVERAGE FOR CHEMICAL ANALYSIS
μ_2 DENOTES THE POPULATION AVERAGE FOR X RAY ANALYSIS

$H_0 : \mu_2 - \mu_1 = 0$ TEST WHETHER OR NOT THERE IS A DIFFERENCE CHEMICAL & XRAY ANALYSIS

$H_1 : \mu_1 - \mu_2 \neq 0$ TEST WHETHER THE TWO METHODS GIVE THE SAME RESULT

$$\underline{H_0 : \overset{\text{X RAY}}{\mu_2} - \underset{\text{CHEMICAL}}{\mu_1} = 0}$$

$d_1 = 2.2 - 2.0 = 0.2$
$d_2 = 1.9 - 2.0 = -0.1$
$d_3 = 2.5 - 2.3 = 0.2$
$d_4 = 2.3 - 2.1 = 0.2$
$d_5 = 2.4 - 2.4 = 0$

$$\bar{d} = \frac{0.2 + -0.1 + 0.2 + 0.2 + 0}{5} = 0.1$$

$\bar{d}$ IS THE SAMPLE MEAN

$$(d_1 - \bar{d})^2 = (0.2 - 0.1)^2 = 0.01$$
$$(d_2 - \bar{d})^2 = (-0.1 - 0.1)^2 = 0.04$$
$$(d_3 - \bar{d})^2 = (0.2 - 0.1)^2 = 0.01$$
$$(d_4 - \bar{d})^2 = (0.2 - 0.1)^2 = 0.01$$
$$(d_5 - \bar{d})^2 = (0 - 0.1)^2 = 0.01$$

TEST STATISTIC

$$t = \frac{\bar{d} - d_0}{S_d/\sqrt{N}} = \frac{0.1 - 0}{S_d/\sqrt{5}}$$

$\bar{d}$ SAMPLE MEAN
d_0 DIFFERENCE BETWEEN THE POPULATION MEANS UNDE NULL HYPOTHESIS WHICH IS 0 IN OUR CASE
S_d SAMPLE STANDARD DEVIATION OF THE DIFFERENCES BETWEEN PAIRED OBSERVATIONS

$$S_d^2 = \frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2 = \left(1/5-1\right)(d_1 - \bar{d})^2 + (d_2 - \bar{d})^2 + (d_3 - \bar{d})^2 + (d_4 - \bar{d})^2 + (d_5 - \bar{d})^2$$
$$= \left(1/4\right)\ (0.2 - 0.1)^2 + (-0.1 - 0.1)^2 + (0.2 - 0.1)^2 + (0.2 - 0.1)^2 + (0 - 0.1)^2$$
$$= \frac{0.01 + 0.04 + 0.01 + 0.01 + 0.01}{4}$$
$$= \frac{0.08}{4}$$
$$S_d^2 = 0.02$$

$$S_d = \sqrt{S_d^2} = \sqrt{0.02} = 0.1414$$

$$t = \frac{\bar{d} - d_0}{S_d/\sqrt{N}} = \frac{0.1 - 0}{S_d/\sqrt{5}} = \frac{0.1 - 0}{0.1414/\sqrt{5}} = 1.58$$

$t = 1.58$
THEREFORE THE CRITICAL REGION IS $(-\infty, -2.7764) \cup (2.7764, \infty)$
SINCE $-2.7764 < 1.58 < 2.7764$ IS TRUE WE CANNOT REJECT THE NULL HYPOTHESIS & CONCLUDE THAT THE TWO METHODS OF ANALYSIS XRAY & CHEMICAL GIVE THE SAME RESULT

probability and statistics morgan bergen

10 one and two sample tests of hypotheses may 1 2023

statistical hypotheses general concepts	tests on a single mean variance unknown
single sample tests concerning a single mean	inferences on paired samples
tests on a single mean variance known	one sample test on a single proportion
relationship to confidence interval estimation	two samples test on two proportions

10.46 in Review Exercise 9.91 on page 313, use the t- distribution to test the hypothesis that the diet reduces a woman's weight by 4.5 kilograms on average against the alternative hypothesis that the mean difference in weight is less than 4.5 kilograms. use a P - value.

9.91 It is claimed that a new diet will reduce a person's weight by 4.5 kilograms on average in a period of 2 weeks. The weights of 7 women who followed this diet were recorded before and after the 2-week period.

woman	weight before	weight after
1	58.5	60.0
2	60.3	54.9
3	61.7	58.1
4	69.0	62.1
5	64.0	58.5
6	62.6	59.9
7	56.7	54.4

Test the claim about the diet by computing a 95% confidence interval for the mean difference in weights. Assume the differences of weights to be approximately normally distributed.

$N = 7$ TOTAL WOMEN

$\alpha = 0.05$ DEFAULT LEVEL OF SIGNIFICANCE

$H_0 : \mu_1 - \mu_2 = 4.5$ NULL HYPOTHESIS

"CLAIMED THAT A NEW DIET REDUCES WEIGHT ON AVERAGE OF 4.5"

$H_1 : \mu_1 - \mu_2 < 4.5$ ALTERNATIVE HYPOTHESIS

" MEAN DIFFERENCE IN WEIGHT IS LESS THAN 4.5"

$d_N = WB_N - WA_N$

$d_1 = 58.5 - 60.0 = -1.5$

$d_2 = 60.3 - 54.9 = 5.4$

$d_3 = 61.7 - 58.1 = 3.6$

$d_4 = 69.0 - 62.1 = 6.9$

$d_5 = 64.0 - 58.5 = 5.5$

$d_6 = 62.6 - 59.9 = 2.7$

$d_7 = 56.7 - 54.4 = 2.3$

$\bar{d}$ SAMPLE MEAN

$$\bar{d} = \frac{-1.5 + 5.4 + 3.6 + 6.9 + 5.5 + 2.7 + 2.3}{7} = 3.55$$

t TEST STATISTIC

$$t = \frac{\bar{d} - d_0}{S_d / \sqrt{N}}$$

$\bar{d}$ SAMPLE MEAN OF THE DIFFERENCES BETWEEN PAIRED OBSERVATIONS

d_0 DIFFERENCE BETWEEN THE POPULATION MEAN UNDER H_0

S_d SAMPLE STANDARD DEVIATION

$(d_1 - \bar{d})^2 = -1.5 - 3.55 = -5.05$

$(d_2 - \bar{d})^2 = 5.4 - 3.55 = 1.85$

$(d_3 - \bar{d})^2 = 3.6 - 3.55 = 0.05$

$(d_4 - \bar{d})^2 = 6.9 - 3.55 = 3.35$

$(d_5 - \bar{d})^2 = 5.5 - 3.55 = 1.95$

$(d_6 - \bar{d})^2 = 2.7 - 3.55 = -0.85$

$(d_7 - \bar{d})^2 = 2.3 - 3.55 = -1.25$

STANDARD DEVIATION OF SAMPLE MEAN

$$S^2_d = \frac{1}{N-1} \sum (d_i - \bar{d})^2 = \frac{1}{7-1} \left(-5.05 + 1.85 + 0.05 + 3.35 + 1.95 + -0.85 + -1.25 \right. = 7.67$$

$S^2_d = 7.67$

$S_d = \sqrt{S^2_d} = 2.77$

TEST STATISTIC

$$t = \frac{\bar{d} - d_0}{S_d / \sqrt{N}} = \frac{3.55 - 4.5}{2.77 / \sqrt{7}} = -0.9$$

6 DEGREES OF FREEDOM

$d_f = N - 1 = 7 - 1 = 6$

$$P_{VALUE} = P(t_{df} \leq t) = P(t_6 \leq -0.9)$$
$$= P(t_{14} \geq 0.9)$$
$$= 0.2$$

PVALUE = 0.2

LEVEL OF SIGNIFICANCE = 0.05

0.2 > 0.05

THEREFORE WE FAIL TO REJECT THE NULL HYPOTHESIS H_0

∴ A DIET REDUCES WOMENS WEIGHT BY 4.5 KILOGRAMS ON AVG

10.57 A new radar device is being considered for a certain missile defense system. The system is checked by experimenting with aircraft in which a kill or a no kill is simulated.

If, in 300 trials, 250 kills occur, accept or reject, at the 0.04 level of significance, the claim that the probability of a kill

with the new system does not exceed the 0.8 probability of the existing device.

GIVEN INFORMATION,

$N = 300$ SAMPLE SIZE
$x = 250$ OCCURENCES SUCCESS
$\alpha = 0.04$ LEVEL OF SIGNIFICANCE

$H_0 : p = 0.8$ NULL HYPOTHESIS

$H_1 : p < 0.8$ ALTERNATIVE HYPOTHESIS

Z VALUE FOR TESTING $P = P_0$ IS

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{N}}}$$ WHICH IS THE VALUE OF THE STANDARD NORMAL VARIABLE Z

IF X REPRESENTS THE NUMBER OF KILLS DURING N TRIALS,
THEN THE PROPORTIONS OF SUCCESSFUL KILLING IS

$$\hat{p} = \frac{x}{N} = \frac{250}{300} = 0.83$$

Z VALUE FOR TESTING $P = 0.8$ IS

$$z = \frac{0.83 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{300}}} \simeq 1.3$$

$Z_{0.04} = 1.75$ OR CRITICAL REGION IS $z < -1.75$
$z = 1.3 > 1.75 = -Z_{0.04}$
THUS WE DO NOT REJECT H_0 & CONCLUDE THAT THE PROBABILITY
OF A KILL WITH A NEW SYSTEM DOES NOT EXCEED THE 0.8 PROBABILITY

10.59 A fuel oil company claims that one-fifth of the homes in a certain city are heated by oil.

Do we have reason to believe that fewer than one-fifth are heated by oil if, in a random sample of 1000 homes in this city, 136 are heated by oil? Use a P-value in your conclusion.

$N = 1000$ SAMPLE SIZE

$x = 136$ NUMBER OF HEATED HOMES

$\alpha = 0.05$ LEVEL OF SIGNIFICANCE

$H_0 : p = 1/5$ NULL HYPOTHESIS

"OIL COMPANY CLAIMS THAT 1/5 OF HOMES ARE HEATED"

$H_1 : p < 1/5$ ALTERNATIVE HYPOTHESIS

"FEWER THAN 1/5 OF HOMES ARE HEATED"

$\hat{p} = x/N$ SAMPLE PROPORTION

$\hat{p} = 136/1000$

$\hat{p} = 0.136$

TEST STATISTIC

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{N}}} = \frac{0.136 - 1/5}{\sqrt{\frac{1/5\,(1 - 1/5)}{1000}}} = \frac{-0.064}{0.0126} = -5.06$$

$P(Z < -5.06) = 0$

$0 < 0.05$, SO WE CAN REJECT THE NULL HYPOTHESIS

& CONCLUDE THAT H_1 IS CORRECT THAT FEWER THAN 1/5 ARE HEATED IN OIL

10.62 In a controlled laboratory experiment, scientists at the University of Minnesota discovered that 25% of a certain strain of rats subjected to a 20% coffee bean diet and then force-fed a powerful cancer-causing chemical later developed cancerous tumors. Would we have reason to believe that the proportion of rats developing tumors when subjected to this diet has increased if the experiment were repeated and 16 of 48 rats developed tumors? Use a 0.05 level of significance.

$n = 48$ SAMPLE SIZE

$H_0 : p = 0.25$ NULL HYPOTHESIS

$H_1 : p > 0.25$ ALTERNATIVE HYPOTHESIS

$\alpha = 0.05$ LEVEL OF SIGNIFICANCE

$x = 16$ TOTAL RATS WITH TUMOR

SAMPLE PROPORTION

$$\hat{p} = x/n = 16/48 = 0.333$$

TEST STATISTIC

$$Z = \frac{0.333 - 0.25}{\sqrt{\frac{0.25(1-0.25)}{48}}} = 1.33 \qquad Z_{0.05} = 1.645$$

$\therefore$ THE # OF RATS THAT HAVE A TUMOR INCREASED