Math 3081, Probability and Statistics, Summer 2, 2019

Test 3

Sample. Solutions

Please **TYPE** your **First Name** and **Last Name**:

Please show all your work, justify your conclusions and give the final answer as a number, either fraction or decimal.

Make the hypothesis testing by performing the following steps.

1. State the null and alternative hypothesis.
2. Write down what data is given in the problem.
3. Find a value of statistic you are going to use.
4. Find critical value/values (or critical region) of your hypothesis.
5. Check is your statistic in a critical region or not.
6. Make a decision about the null and alternative hypothesis.
7. Answer the question of the problem.

In case when other parameters are required to find in the problem please also include them in your answer.

Problem 1.

Suppose you want to test $H_0 : \mu = 12$ versus $H_1 : \mu > 12$ at significance 0.05. The sample mean is 13.2, the size of the sample is 25. Will you reject the null hypothesis if

(a) (10 pts) sample standard deviation is 2.5?

(b) (10 pts) standard deviation of distribution is 2.5?

$\sigma = 2.5$

(a) $H_0 : \mu = 12$

$H_1 : \mu > 12$

$\alpha = 0.05$

$\bar{x} = 13.2$

$n = 25 \Rightarrow d.f. = 24$

$s = 2.5$

$$t = \frac{\bar{x} - 12}{s/\sqrt{n}} = 2.4$$

$t_{\alpha; d.f.} = t_{0.05; 24} = 1.711$

$t > t_{\alpha; d.f.} \Rightarrow$ $\underline{H_0 \text{ - reject}}$

(P-value $= 0.012 < \alpha$)

(b) $$z = \frac{\bar{x} - 12}{\sigma/\sqrt{n}} = 2.4.$$

$z_\alpha = 1.65$

$z > z_\alpha \Rightarrow$ $\underline{H_0 \text{ - reject}}$

(P-value $= 0.008 < \alpha$)

Problem 2. (5 pts)

If a normally distributed sample of size 16 produces a 95% confidence interval for a mean μ that ranges from 44.7 to 49.9, what are the values of $\overline{x}$ and s? (assume that the **variance is unknown**)

$n = 16 \Rightarrow d.f. = 15$

$\alpha = 0.05 \Rightarrow t_{\frac{\alpha}{2}; d.f.} = t_{0.025; 15} = 2.131$

$\left(\bar{x} - t_{\frac{\alpha}{2}; d.f.} \cdot \frac{s}{\sqrt{n}} ; \bar{x} + t_{\frac{\alpha}{2}; d.f.} \frac{s}{\sqrt{n}}\right) = (44.7 ; 49.9)$

$\bar{x} = \frac{44.7 + 49.9}{2} = \underline{\underline{47.3}}$

$2.131 \cdot \frac{s}{\sqrt{16}} = 49.9 - 47.3$

$s = \frac{2.6 \cdot 4}{2.131} = \underline{\underline{4.88}}$

Problem 3. (10 pts)

A worker, looking for a new position, wonders if the salary difference between type X and Y institutions is really significant. He finds that a random sample of 200 type X institutions has a mean salary of 54 thousands, with a standard deviation $\sigma_X = 8.5$ thousands. A random sample of 200 type Y institutions has a mean salary of 46 thousands, with a standard deviation $\sigma_Y = 8.1$ thousands. Do these data indicate a **significantly higher** salary at type X institutions?

$n = 200$ $\quad m = 200$

$\bar{x} = 54$ $\quad \bar{y} = 46$

$\sigma_x = 8.5$ $\quad \sigma_y = 8.1$

$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} = 9.6359 > Z_\alpha = 1.65$

So, $\underline{\text{Reject } H_0}$

$H_0: \mu_x = \mu_y$ $\quad \alpha = 0.05$

$H_1: \mu_x > \mu_y$

$p\text{-value} = 2.87 \cdot 10^{-22} < \alpha$

Problem 4. (5 pts)

A professor wants to compare the students' scores with the national average. The professor chooses 20 students, who score an average $\overline{x} = 50.2$ on a standardized test. Their scores have a sample standard deviation of $s = 2.5$. The national average on the test is a 60. The professor wants to know are the obtained scores **significantly lower** than the national average or not.

$n = 20 \Rightarrow d.f. = 19$

$\bar{x} = 50.2$

$\mu_0 = 60$ $\quad \alpha = 0.05$

$s = 2.5$

$t_{\alpha; d.f.} = t_{0.05; 19} = 1.729$

$H_0: \mu = 60$

$H_1: \mu < 60$

$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = -17.53$

$t < -t_{\alpha; d.f.} \Rightarrow \underline{\text{Reject } H_0}$

$(p\text{-value} \approx 0 < \alpha)$

Problem 5. (5 pts)

The grades in a test are known to follow a normal distribution, but the mean and variance are unknown. A random sample of 6 students produced the following scores: 8, 33, 22, 12, 5, 10. Use the t-distribution to find a 90% two-sided confidence interval for the mean score.

$n = 6$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
8	-7	49
33	18	324
22	7	49
12	-3	9
5	-10	100
10	-5	25

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$\bar{x} = 15$

$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 556/5 = 111.2$

$s = 10.545$

$\alpha = 0.1 \Rightarrow t_{\frac{\alpha}{2}; d.f.} = t_{0.05; 5} = 2.015$

$d.f. = 5$

$CI: \bar{x} \pm t_{\frac{\alpha}{2}; d.f.} \cdot \frac{s}{\sqrt{n}} = 15 \pm 2.015 \cdot \frac{10.545}{\sqrt{6}} =$

$= 15 \pm 8.67 =$

$= \underline{\underline{(6.33; 23.67)}}$

Problem 6.

A random sample of 10 male newborns has a mean weight of 123 ounces and a sample standard deviation $s_X = 8$ ounces, while a sample of 8 female has a mean weight of 116 ounces and a sample standard deviation of $s_Y = 5$ ounces. Assume that the **population variances are the same.**

(i) (10 pts) Test at 0.05 level of significance if there is a **significant difference** between the weights of the two groups.

(ii) (10 pts) Find the 95% confidence interval for the difference between the means.

$n=10$ $\quad m=8$

$\bar{x}=123$ $\quad \bar{y}=116$

$s_X=8$ $\quad s_Y=5$ $\quad \boxed{\text{Pooled}}$

$\alpha=0.05;\ \sigma_X=\sigma_Y$

$d.f.=n+m-2=16$

$$\begin{cases} H_0: \mu_X=\mu_Y \\ H_1: \mu_X\neq\mu_Y \end{cases}$$

$t_{\frac{\alpha}{2};d.f.}=t_{0.025;16}=2.12$

(a) $t=\dfrac{\bar{x}-\bar{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}}=\dfrac{123-116}{6.85\sqrt{\frac{1}{10}+\frac{1}{8}}}=2.1543$

$s_p^2=\dfrac{(n-1)s_X^2+(m-1)s_Y^2}{n+m-2}=\dfrac{9\cdot 64+7\cdot 25}{16}=$

$=49.9375$

$s_p=6.85$

$|t|>t_{\frac{\alpha}{2};d.f.}\Rightarrow$ Reject H_0

(P-value $=0.047$)

(b) CI: $\bar{x}-\bar{y}\pm t_{\frac{\alpha}{2};d.f.}\cdot s_p\sqrt{\frac{1}{n}+\frac{1}{m}}=7\pm 6.89=(0.11;\ 13.89)$

Problem 7. (10 pts)

Given the width-to-length ratios with $\overline{x}=0.637$ and $s=0.141$ for a random sample of the flags of 34 countries check the hypothesis that the width-to-length ratio for national flags is 0.618 with 0.05 level of significance. Assume that the alternative hypothesis is two-sided. Additionally, find a 90% confidence interval for a mean.

$\bar{x}=0.637$

$s=0.141$

$n=34\Rightarrow d.f.=33$

$\mu_0=0.618$

$\alpha=0.05$

$$\begin{cases} H_0: \mu=0.618 \\ H_1: \mu\neq 0.618 \end{cases}$$

$t_{\frac{\alpha}{2};d.f.}=t_{0.025;33}=2.04$

$t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}=0.7857$ $\quad$ (P value $=0.438$)

$|t|\ngtr t_{\frac{\alpha}{2};d.f.}\Rightarrow$ Accept H_0

90% CI: $\bar{x}\pm t_{\frac{\alpha}{2};d.f.}\dfrac{s}{\sqrt{n}}=$

$\Downarrow$

$\alpha=0.1$

$=0.637\pm 1.69\cdot\dfrac{0.141}{\sqrt{34}}=$

$=0.637\pm 0.041=$

$=(0.596;0.678)$

Problem 8. (5 pts)

The manufacturer's tests have shown that in a year's time in the particular kind of soil the manufacturer must deal with, the average depth of the maximum pit in a foot of pipe is $\mu_0 = 0.0042$ inch. To see whether that average can be reduced, ten pipes are coated with a new plastic and buried in the same soil. After one year, the average maximum pit depth is recorded being 0.0039 inch. Given that the sample standard deviation for these 10 measurements is $s = 0.00383$ inch, can it be concluded at the 0.05 level of significance that the plastic coating is beneficial (i.e. is the **average reduced** or the same)?

$\mu_0 = 0.0042$

$n = 10 \Rightarrow d.f. = 9$

$s = 0.00383$

$\bar{x} = 0.0039$

$\alpha = 0.05$

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$$

$t_{\alpha; d.f.} = t_{0.05; 9} = 1.833$

$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = -0.2477$

$t \nless -t_{\alpha; d.f} \Rightarrow$

$\Rightarrow$ Accept H_0

(P-value $= 0.40496$)

Problem 9.

A medical researcher believes that women typically have lower serum cholesterol than men. To test this hypothesis, he took a sample of 476 men and found their mean serum cholesterol to be 189 mg/dl with a sample standard deviation $s_X = 34.2$. A group of 592 women had an average 177.2 mg/dl and a sample standard deviation $s_Y = 33.3$. Set 0.05 level of significance and use **pooled variance**.

(a) (5 pts) Is the **lower average** for the women statistically significant?

(b) (5 pts) Find 95% confidence interval for the difference of means.

Pooled

a) $n = 476$ $\quad m = 592$

$\bar{x} = 189$ $\quad \bar{y} = 177.2$

$s_X = 34.2$ $\quad s_Y = 33.3$

$\alpha = 0.05$. $\quad \sigma_X = \sigma_Y$

$$\begin{cases} H_0: \mu_X = \mu_Y \\ H_1: \mu_X > \mu_Y \end{cases}$$

$d.f. = n + m - 2 = 1066$

$t_{\alpha; d.f.} \approx z_\alpha = 1.65$

$s_p^2 = \dfrac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2} = 1135.96$

(P-value $= 8.34 \cdot 10^{-9}$)

$s_p = 33.7$

$t = \dfrac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} = 5.688 > t_{\alpha; d.f} \Rightarrow$

$\Rightarrow$ Reject H_0

b) 95% CI: $(\bar{x} - \bar{y}) \pm t_{\frac{\alpha}{2}; d.f.} \cdot s_p \sqrt{\frac{1}{n} + \frac{1}{m}} =$

$= 11.8 \pm 4.066 = (7.734 \,;\, 15.866)$

Problem 10. (10 pts)

In a comparison of these two diets for one-year weight loss, a study found that 77 subjects on the diet A had an average weight loss of 4.7 kg and a sample standard deviation $s_X = 7.05$ kg. Similar figures for the 79 people on the diet B were 1.6 kg and $s_Y = 5.36$ kg. Is the **greater reduction** with the diet A statistically significant? Test for 0.05 level of significance (use **non-pooled variance**).

$n = 77$ $\quad m = 79$

$\bar{x} = 4.7$ $\quad \bar{y} = 1.6$

$s_X = 7.05$ $\quad s_Y = 5.36$

$\alpha = 0.05$; $\quad \sigma_X \neq \sigma_Y$ – non-pooled

$$\begin{cases} H_0: \mu_X = \mu_Y \\ H_1: \mu_X > \mu_Y \end{cases}$$

$t = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m}}} = 3.086$

$t_{\alpha; d.f.} \approx t_{0.05; 142} = 1.66$

$t > t_{\alpha; d.f.} \Rightarrow$ Reject H_0

$d.f. = 141.88$

(P-value $= 0.0012 < \alpha$)