STAT 403 Tutorial · Week 3

For this week's tutorial, I will give two numerical examples one-sample and two-sample problems, including hypothesis testing and confidence intervals. In addition, more details about the randomization test introduced in class will also be provided.

(I). One-sample problem

Example 1. Marks on an exam in a statistics course are assumed to be normally distributed with some unknown mean μ_1 and unknown variance σ. You can think of μ_1 as the center location (or average) of all marks. Suppose that someone claim that $\mu_1 = 60$. To evaluate whether this statement is true. A random sample of 4 students is selected, and their marks are 52, 63, 64, 84. Assess the statement and compute a 95% confidence interval for μ_1.

Analysis: To make it concise, we use X_1, X_2, X_3, X_4 to denote the four marks in sample. Here the sample size $n=4$. By assumptions, $X_1, \dots, X_4 \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$. One may expect that, if $\mu_1 = 60$ is true, the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{4} X_i = \frac{1}{4}(X_1+X_2+X_3+X_4)$ should be close to 60; in other words, $\bar{X} - 60$ should be close 0. Here in our example, $\bar{X} = \frac{1}{4}(52+63+64+84) = 65.75 \Rightarrow \bar{X} - 60 = 5.75$. But whether this 5.75 is close enough to 0? To give a more formal answer, we need to use some statistical theory. Technically, both hypothesis testing and confidence interval are based on the fact that if $X_1, \dots, X_n \overset{iid}{\sim} N(\mu_1, \sigma^2)$, then $\frac{\bar{X} - \mu_1}{\sqrt{S_1^2/n}} \sim t_{n-1}$, where $S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Solution: 1. hypothesis testing

Suppose $H_0: \mu = 60$ is true, then $T = \frac{\bar{X} - 60}{\sqrt{S_1^2/n}}$ is supposed to have a t distribution with $n-1 = 4-1 = 3$ degrees of freedom. Here,

$$\bar{X} = \frac{1}{4}(52+63+64+84) = 65.75 \; ; \quad S_1^2 = \frac{1}{4-1}\left[(52-65.75)^2 + (63-65.75)^2 + (64-65.75)^2 + (84-65.75\right.$$
$$= 177.58$$

Therefore, the observed value for T is $T_{obs} = \frac{65.75 - 60}{\sqrt{177.58/4}} = 0.86$

Now we are ready to calculate p-values. Definition of p-values varies according to the choice of alternative hypothesis.

① if $H_a: \mu \neq 60$, then p-value $= \mathbb{P}(|T| \geq |T_{obs}|) = \mathbb{P}(|T| \geq 0.86) = \mathbb{P}(T \geq 0.86) + \mathbb{P}(T \leq -0.86) \approx 0.45$;

② if $H_a: \mu > 60$, then p-value $= \mathbb{P}(T \geq T_{obs}) = \mathbb{P}(T \geq 0.86) \approx 0.23$;

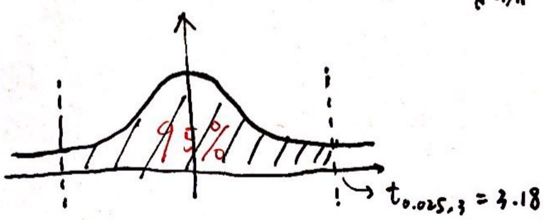
③ if $H_a: \mu < 60$, then p-value $= \mathbb{P}(T \leq T_{obs}) = \mathbb{P}(T \leq 0.86) \approx 0.77$.

If we set the significance level to be 0.05, then we reject the null hypothesis if p-values are less than 0.05. Since all three p-values are above 0.05, we can't reject $H_0: \mu = 60$ and accept any of the three hypotheses.

2. Confidence interval

Now instead of assessing whether a statement about μ_1 is true or false, we try to give a range of values (interval) such that we are 95% certain that μ_1 is contained in this interval.

Since $\dfrac{\bar{X} - \mu_1}{\sqrt{S_1^2/n}} \sim t_{n-1}$, we have $\mathbb{P}\left(\left| \dfrac{\bar{X} - \mu_1}{\sqrt{S_1^2/n}} \right| \leq t_{0.025,4-1} \right) = 95\%$



Note that $\left| \dfrac{\bar{X} - \mu_1}{\sqrt{S_1^2/n}} \right| \leq t_{0.025,3} \iff \bar{X} - t_{0.025,3}\sqrt{\dfrac{S_1^2}{n}} \leq \mu \leq \bar{X} + t_{0.025,3}\sqrt{\dfrac{S_1^2}{n}}$

$$\Rightarrow \mathbb{P}\left(\mu \in \left[\bar{X} - t_{0.025,3}\sqrt{\frac{S_1^2}{n}},\ \bar{X} + t_{0.025,3}\sqrt{\frac{S_1^2}{n}} \right] \right) = 95\%$$

By plugging in $\bar{x} = 65.75$, $S_1^2 = 177.58$, $n = 4$, $t_{0.025,3} = 3.18$, we obtain a 95% confidence interval for μ: $\left[\bar{x} - t_{0.025,3}\sqrt{\frac{S_1^2}{n}},\ \bar{x} + t_{0.025,3}\sqrt{\frac{S_1^2}{n}} \right] = [44.56, 86.94]$.

(II). Two-sample problem

Example 2 (Example 1 Ctd.) Suppose that a new teaching method is applied to another class of students. The same exam was taken by these students and the marks are assumed to be normally distributed with mean μ_2 and same unknown variance. The administrator wants to evaluate whether the new teaching method has better performance by assessing whether $\mu_2 > \mu_1$. In addition to sample collected in Example 1, a sample of 5 students taught by new teaching method is collected, and their marks are 51, 58, 66, 77, 80. Assess the hypothesis $H_0: \mu_1 = \mu_2$ against $H_a: \mu_2 > \mu_1$ and determine a 95% CI for $\mu_2 - \mu_1$.

Analysis: The question can be summarized as follows,

$$X_1, X_2, X_3, X_4 \overset{iid}{\sim} N(\mu_1, \sigma^2), \quad \text{sample size } n=4;$$
$$Y_1, Y_2, Y_3, Y_4, Y_5 \overset{iid}{\sim} N(\mu_2, \sigma^2), \quad \text{sample size } m=5.$$

Similarly, we can use $\bar{X} = \frac{1}{4}(X_1+X_2+X_3+X_4) = 65.75$ to estimate μ_1,

$\bar{Y} = \frac{1}{5}(Y_1+Y_2+Y_3+Y_4+Y_5) = 66.4$ to estimate μ_2.

And if μ_1 and μ_2 are close, $\bar{x}$ and $\bar{y}$ should be close. Again, to make it strict, we need to use the fact that

$$T = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{S^2(\frac{1}{n} + \frac{1}{m})}} \sim t_{n+m-2}, \quad \text{where } S^2 = \frac{1}{n+m-2}\left[(n-1)S_1^2 + (m-1)S_2^2\right].$$
$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$
$$S_2^2 = \frac{1}{m-1}\sum_{j=1}^{5}(Y_j - \bar{Y})^2.$$

Solution:

1. hypothesis testing

Under the null hypothesis $H_0: \mu_1 = \mu_2$, T becomes $T = \frac{(\bar{Y} - \bar{X}) - 0}{\sqrt{S^2(\frac{1}{n} + \frac{1}{m})}}$.

One may calculate the observed value of T: $T_{obs} = \frac{66.4 - 65.75}{\sqrt{162.56(\frac{1}{4} + \frac{1}{5})}} = 0.08$,

where $162.56 = S^2 = \frac{1}{4+5-2}\left[(4-1)S_1^2 + (5-1)S_2^2\right]$, $S_1^2 = 177.58$, $S_2^2 = \frac{1}{5-1}\left[(51-66.4)^2 + \dots\right] = 151$

The p-value is calculated as p-value $= \mathbb{P}(T > T_{obs}) = 0.47 > 0.05$.

Therefore we do not find sufficient evidence against $\mu_1 = \mu_2$ and cannot reject the null hypothesis.

2. Confidence interval.

By $\mathbb{P}(|T| \le t_{0.025,7}) = 95\%$ and plugging in expression of T,

the 95% C.I. is $(\bar{Y} - \bar{X}) \pm t_{0.025,7}\sqrt{S^2(\frac{1}{n}+\frac{1}{m})} =$ [redacted]. [-19.54, 20.84]

(III). Randomization test.

Recall: we are asked to compare two fertilizers A and B to see if B is better than A. 3 A's and 3 B's are randomly assigned to 6 plots.

	1	2	3	4	5	6
	A	B	B	A	A	B
Y:	11	12	16	13	9	14.

We use a measure of difference $D = \bar{Y}_B - \bar{Y}_A = \frac{1}{3}(12+16+14) - \frac{1}{3}(11+13+9) = 3$

If A and B are the same, then unit 1 will produce the same yield whether treatment A or B is applied. So if another set of treatments are given, for example,

	1	2	3	4	5	6
	B	A	A	A	B	B
Y	11	12	16	13	9	14

→ does not change

Then, $D = \bar{Y}_B - \bar{Y}_A$
$= \frac{1}{3}(11+9+14) - \frac{1}{3}(12+16+14) = -2.3$

So if A and B are the same, we are equally likely to observe $D=3$ and $D=-2.33$.

Besides, there are other 18 possibilities, because number of possible randomizations is $\binom{6}{3} = 20$:

Treatment assignment	$\bar{Y}_A$	$\bar{Y}_B$	$D = \bar{Y}_B - \bar{Y}_A$	
A A A B B B	13	12	-1	
A A B A B B	12	13	1	
A A B B A B	10.67	14.33	3.67	✓ (because >3).
A A B B B A	12.33	12.67	0.33	
A B A A B B	13.33	11.67	-1.67	
A B A B A B	12	13	1	
A B A B B A	13.67	11.33	-2.33	
A B B A A B	11	14	3	→ observed
A B B A B A	12.67	12.33	-0.33	
A B B B A A	11.33	13.67	2.33	
B A A A B B	13.67	11.33	-2.33	
B A A B A B	12.33	12.67	0.33	
B A A B B A	14	11	-3	
B A B A A B	11.33	13.67	2.33	
B A B A B A	13	12	-1	
B A B B A A	11.67	13.33	1.67	
B B A A A B	12.67	12.33	-0.33	
B B A A B A	14.33	10.67	-3.67	
B B A B A A	13	12	-1	
B B B A A A	12	13	1	

Each of 20 D's observed with probability 1/20.

Therefore the p-value $= \mathbb{P}(D > D_{obs}) + \frac{1}{2}\mathbb{P}(D = D_{obs})$

$= \frac{1}{20} + \frac{1}{2} \times \frac{1}{20}$

$= \frac{3}{40} = 0.075.$