生物統計 HW8 解答

#1:

Red wine is good for the heart. Observational studies suggest that moderate use of alcohol reduces heart attacks, and that red wine may have special benefits. One reason may be that red wine contains polyphenols, substances that do good things to cholesterol in the blood and so may reduce the risk of heart attacks. In an experiment, healthy men were assigned at random to several groups. One group of 9 men drank half a bottle of red wine each day for two weeks. The level of polyphenols in their blood was measured before and after the two-week period. Here are the percent changes in level:[2]

3.5 8.1 7.4 4.0 0.7 4.9 8.4 7.0 5.5

Make a stemplot of the data. It is difficult to assess Normality from just observations. Give a 90% t confidence interval for the mean percent change in blood polyphenols among all healthy men if all drank this amount of red wine.

Sol):

Stemplot: (自己畫，check 是否服從常態)

$$\bar{x} = 5.5, s = 2.5169, t_{8,0.05} = 1.8595$$

$$90\%\ C.I.\text{ of }\mu = [5.5 - 1.8595 \times \frac{2.5169}{\sqrt{9}}, 5.5 + 1.8595 \times \frac{2.5169}{\sqrt{9}}] = [3.9399, 7.0601]$$

#2:

Ancient air. The composition of the earth's atmosphere may have changed over time. To try to discover the nature of the atmosphere long ago, we can examine the gas in bubbles inside ancient amber. Amber is tree resin that has hardened and been trapped in rocks. The gas in bubbles within amber should be a sample of the atmosphere at the time the amber was formed. Measurements on specimens of amber from the late Cretaceous era (75 to 95 million years ago) give these percents of nitrogen:[3]

63.4 65.0 64.4 63.3 54.8 64.5 60.8 49.1 51.0

Assume (this is not yet agreed on by experts) that these observations are an SRS from the late Cretaceous atmosphere.

(a)

Graph the data, and comment on skewness and outliers. The t procedures will be only approximate for these data.

Sol):

畫完自行判別。

(b)

Give a 95% t confidence interval for the mean percent of nitrogen in ancient air.

Sol):

$\bar{x} = 59.5889, s = 6.2553, t_{8,0.025} = 2.3060$

95% $C.I.$ of $\mu = [\bar{x} - 2.3060 \times \frac{s}{\sqrt{9}}, \bar{x} + 2.3060 \times \frac{s}{\sqrt{9}}] = [54.7807, 64.3971]$

#3:

Student study times. A student group claims that first-year students at a university must study 2.5 hours per night during the school week. A skeptic suspects that they study less than that on the average. A class survey finds that the average study time claimed by 269 students is $\bar{x} = 137$ minutes. Regard these students as a random sample of all first-year students and suppose we know that study times follow a Normal distribution with standard deviation 65 minutes. Carry out a test of H_0: $\mu = 150$ against H_a: $\mu < 150$. What do you conclude?

Sol):

$H_0 : \mu = 150$ v.s. $H_1 : \mu < 150$

$\overline{X} = 137, n = 269, \sigma = 65$

$if \;\; \alpha = 0.05$

檢定統計量: $Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$

檢定規則: $reject \; H_0 \; if \; Z < -Z_{0.05} = -1.645$

$\because z = \frac{137 - 150}{65 / \sqrt{269}} = -3.2802 < -1.645$

$\therefore reject \; H_0 \; at \; \alpha = 0.05 \; level$

表示有足夠證據顯示學生平均念書時間少於 150 分鐘

#4:

IQ test scores. Exercise 13.6 (page 329) gives the IQ test scores of 31 seventh-grade girls in a Midwest school district. IQ scores follow a Normal distribution with standard deviation $\sigma = 15$. Treat these 31 girls as an SRS of all seventh-grade girls in this district. IQ scores in a broad population are supposed to have mean $\mu = 100$. Is there evidence that the mean in this district differs from 100? State hypotheses, find the test statistic and its P-value, and state your conclusion.

Sol):

$H_0 : \mu = 100$ v.s. $H_1 : \mu \neq 100$

$\overline{X} = 105.83, n = 31, \sigma = 15$

$if \ \ \alpha = 0.05$

檢定統計量: $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$

檢定規則: $reject\ H_0\ \ if\ \ |Z| > Z_{0.025} = 1.96$

$\because |z| = \left| \dfrac{105.83 - 100}{15/\sqrt{31}} \right| = 2.164 > 1.96$

$\therefore reject\ H_0\ \ at\ \ \alpha = 0.05\ level$

表示有足夠證據顯示 IQ 平均分數不等於 100

#5:

請針對 (3) (4) 分別導出 rejection rule

參考 Lecture note 決策方法 B

Sol):

(a) 第三題:

檢定規則:

1. $reject\ H_0\ \ if\ \ Z < -Z_{0.05} = -1.645$

$\because z_{obs} = \dfrac{137 - 150}{65/\sqrt{269}} = -3.2802 < -1.645$

$\therefore reject\ H_0\ \ at\ \ \alpha = 0.05\ level$

2. $reject\ H_0\ if\ \ p-value = \Pr(Z < z_{obs}) < 0.05$

$\because\ p-value = \Pr(Z < -3.2802) \approx 0.0005 < 0.05$

$\therefore reject\ H_0\ at\ \ \alpha = 0.05\ level$

表示有足夠證據顯示學生平均念書時間少於 150 分鐘

(b) 第四題:

檢定規則:

1. $reject\ H_0\ if\ \ |Z| > Z_{0.025} = 1.96$

$$\because |z| = \left|\frac{105.83-100}{15/\sqrt{31}}\right| = 2.164 > 1.96$$

$\therefore reject\ H_0\ at\ \ \alpha = 0.05\ level$

2. $reject\ H_0\ if\ \ p-value = \Pr(|Z| > |z_{obs}|) < 0.05$

$\because\ p-value = \Pr(|Z| > 2.164) \approx 2 \times 0.0152 = 0.0304 < 0.05$

$\therefore reject\ H_0\ at\ \ \alpha = 0.05\ level$

表示有足夠證據顯示 IQ 平均分數不等於 100

#6:

What's the P-value? A test of the null hypothesis H_0: $\mu = 0$ give test statistic $z = 1.8$.

(a)

What is the P-value if the alternative is H_a: $\mu > 0$?

Sol):

$\Pr(Z > 1.8) = 1 - 0.9641 = 0.0359$

(b)

What is the P-value if the alternative is H_a: $\mu < 0$?

Sol):

$\Pr(Z < 1.8) = 0.9641$

(c)

What is the P-value if the alternative is H_a: $\mu \neq 0$?

Sol):

$2 \times \Pr(Z > 1.8) = 2 \times 0.0359 = 0.0718$

#7:

P and significance. The *P*-value for a significance test is 0.078.

(a)

Do you reject the null hypothesis at level $\alpha = 0.05$? Explain answer.

Sol):

$\because\ p - value = 0.078 > 0.05$

$\therefore do\ not\ reject\ H_0\ at\ \alpha = 0.05\ level$

(b)

Do you reject the null hypothesis at level $\alpha = 0.01$? Explain answer.

Sol):

$\because\ p - value = 0.078 > 0.01$

$\therefore do\ not\ reject\ H_0\ at\ \alpha = 0.05\ level$