Problem Set 3 Solutions

ECON 210 Econometrics A

Evan Zuofu Liao*

October 15, 2015

Question 1

(i) $H_0 : \mu = 0$
(ii) $H_a : \mu < 0$
(iii) The t-statistic is simply

$$\frac{\bar{y} - 0}{s/\sqrt{n}} = -2.11$$

Thus the p-value is

$$\text{p-value} = \Pr(Z \leq -2.11) = .0174$$

So we may reject at the 5% level but not at the 1% level.
(iv) In practice, this question is completely subjective at the personal level. For a moderate drinker, a reduction of 33 ounces in annual consumption does not seem large. On the other hand, when aggregated across the entire population, alcohol producers and distributors might not think the effect is so small.
(v) We assume that other factors that affect liquor consumption, for instance, income and transportation costs etc., are constant for the two years we used.

*Comments and questions to evanliao@uchicago.edu. This solution draws from answers provided by previous TAs.

Question 2

(i) The confidence interval is given by

$$\left[\overline{D}_n - c\sqrt{\frac{s^2}{n}}, \overline{D}_n + c\sqrt{\frac{s^2}{n}}\right]$$

We find that $\overline{D}_n = .24$. The 97.5th percentile in the t_{14} distribution is 2.145. So the confidence interval is

$$.24 \pm 2.145(.1164) = [-.01,\ .49]$$

(ii) $H_0 : \mu = 0$ and $H_a : \mu > 0$
(iii) The t-statistics is $\overline{D}_n/(s/\sqrt{n}) = .24/.1164 \approx 2.062$. The 5% critical value for a one-tailed test with 14 degree of freedom is 1.761, while the 1% critical value is 2.624. Therefore, it is rejected at the 5% level but not the 1% level.
(iv) The p-value obtained from R is .029.

Question 3

(a) The sample analogue is

$$q_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

(b) We can directly apply the Central Limit Theory and find

$$q_n \xrightarrow{d} \mathcal{N}\left(q, \frac{q(1-q)}{n}\right)$$

(c) We can first normalize to get

$$\frac{q_n - q}{\sqrt{\frac{q(1-q)}{n}}} \xrightarrow{d} \mathcal{N}(0,1)$$

Now replace q in the denominator with q_n and apply the Slutsky theorem we know that the resulting statistic can be approximated by a standard normal distribution

$$T_n = \frac{q_n - q}{\sqrt{\frac{q_n(1-q_n)}{n}}} \sim \mathcal{N}(0,1)$$

(d) Remember $Z \sim \mathcal{N}(0,1)$ and $Z_{1-\alpha}$ is defined such that $\Pr(Z \geq Z_{1-\alpha}) = \alpha$. Since T_n is approximately standard normal, replace Z with T_n will give us the desired results. (Note that more formally, we have $\lim_{n\to\infty} \sup \Pr(T_n \geq Z_{1-\alpha}) = \alpha$)

(e) From the previous results we know that the test statistic has to satisfy

$$-Z_{1-\frac{\alpha}{2}} \leq T_n \leq Z_{1-\frac{\alpha}{2}}$$

So a confidence interval of level α is

$$\left[q_n - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{q_n(1-q_n)}{n}}, q_n + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{q_n(1-q_n)}{n}}\right]$$

(f) The p-value is the smallest value of α for which we reject the null, assuming the null hypothesis is true. So we have

$$\text{p-value} = \Pr(|Z| \geq |T_n|) = 2(1-\Phi(T_n))$$

(g) Under the null hypothesis, we have

$$\begin{aligned}
\Pr(T_n \geq Z_{1-\alpha}) = \Pr\left(\frac{q_n}{\sqrt{\frac{q_n(1-q_n)}{n}}} \geq Z_{1-\alpha}\right) &= \Pr\left(\frac{q_n - q}{\sqrt{\frac{q_n(1-q_n)}{n}}} + \frac{q}{\sqrt{\frac{q_n(1-q_n)}{n}}} \geq Z_{1-\alpha}\right) \\
&\leq \Pr\left(\frac{q_n - q}{\sqrt{\frac{q_n(1-q_n)}{n}}} \geq Z_{1-\alpha}\right) = \alpha
\end{aligned}$$

By duality, the confidence interval is the non-rejection region, so we know

$$\begin{aligned}
-\infty < T_n &\leq Z_{1-\alpha} \\
\Rightarrow q_n - Z_{1-\alpha}\sqrt{\frac{q_n(1-q_n)}{n}} &\leq q < \infty
\end{aligned}$$

The p-value for a one-sided test is

$$\text{p-value} = \Pr(Z > T_n) = 1 - \Phi(T_n)$$

Question 4

(a)-(c) Please see the code on Chalk. A few comments here: $\widetilde{\mu} \approx 0.105$, and I cannot reject the hypothesis that $\mu = 0$ at the 5% level. The significance level at which $\widetilde{\mu}$ rejects the null is

15.8%. I find about 65.4% of the new $\hat{\mu}$'s fall above the old 95th percentile. This proportion is called the power of the test.

(d) The standard error is given by

$$SE = \frac{\sqrt{\text{Var}\,(X)}}{\sqrt{n}} = \frac{\sqrt{1/3(a^2 + ab + b^2)}}{\sqrt{n}}$$
$$= \frac{\sqrt{1/3((-0.5)^2 + (-0.5)(0.5) + 0.5^2)}}{\sqrt{10}} \approx 0.0912871$$

Using the CLT, we know that the sample mean is approximately distributed as

$$\hat{\mu} \sim N(0, SE^2)$$

Then the critical value for $\alpha = 0.05$ is $C_\alpha = 0.1501539$

(e) Under $\mu = .2$, the asymptotic approximation becomes

$$\hat{\mu} \sim N(0.2, \widehat{SE}^2)$$

where

$$\widehat{SE} = \frac{\sqrt{\text{Var}\,(X)}}{\sqrt{n}} = \frac{\sqrt{\mathbb{E}\,[X^2] - \mathbb{E}\,[X]^2}}{\sqrt{n}} = \frac{\sqrt{1/3((-0.3)^2 + (-0.3)(0.7) + 0.7^2) - 0.2^2}}{\sqrt{10}} \approx 0.0912871$$

Note that $\widehat{SE}$ should be equal to SE because there is only a mean shift. So the power of the test is given by

$$\Pr(\hat{\mu} > 0.1501539) = 0.7074796 \approx 70.7\%$$

(f) As we can see, the power of the test is slightly higher when using the asymptotic approximation. This is probably because we have a relatively small sample size (remember n=10) so the sampling distribution obtained from simulation is not perfectly normal.