Stat 101
Summer 2019
Dr. Lily Yen

Midterm 3

Show all your work

Name: ______________
Number: ______________
Signature: ______________
Score: ___/39

A TI-83/84 calculator allowed.

Problem 1: Determine the critical values for a 92 % confidence interval from a standard normal distribution.

$$(\text{InvNorm}(0.04), \text{InvNorm}(0.96)) \approx (-1.75, 1.75)$$

Score: /1

Problem 2: In an extensive survey, 76 % of Albertans favour the construction of a pipeline through BC. Assuming that this percentage is true, answer the following.

a. In a random sample of 1500 Albertans, find the 95 % confidence interval estimate of the population proportion p.

   Here $\hat{p} = 0.76$ and $n = 1500$, np and nq both greater than 5, and for 95 % confidence, $z = \text{InvNorm}(0.975) \approx 1.96$, so the endpoints of the confidence interval are $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 0.738, 0.782$

b. How many would we need to sample to be 98 % confident that the sample proportion is within 2 percentage points of the true population proportion?

   Here $z = \text{InvNorm}(0.99) \approx 2.33$, so if $z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.02$, then $n \geq \frac{z^2\hat{p}(1-\hat{p})}{0.02^2} \approx 2467.8$, so we would need to sample at least 2468 Albertans.

Score: /6

Problem 3: To estimate the average number of spam e-mails each CapU student receives in a week, we randomly sample 40 students on campus. Suppose it is known that the standard deviation of this number is 8.7.

a. Find the error if we were to construct a 99 % confidence interval.

   Since the standard deviation is known, you should use the normal distribution. Therefore $z = \text{InvNorm}(0.995) \approx 2.58$. The error is then $z\sigma/\sqrt{n} \approx 3.54$.

b. Find the size of a sample needed to be 99 % confident that the sample mean is within 0.5 spam e-mails of the population mean.

   If $z\sigma/\sqrt{n} = 0.5$, then $n \geq (z\sigma/0.5)^2 \approx 2009$.

Score: /4

/11

Problem 4: Give the type I and type II errors that correspond with the hypothesis: "Watching another movie on Netflix won't put me over my internet download allowance." State first the null hypothesis and the alternative hypothesis.

H_0: Watching another movie yields less than or equal to my internet download allowance.
H_1: Watching another movie puts me over my internet download allowance.
Type I: H_0 is true, but we reject H_0. Claiming that another movie will put you over when you actually have room for more movies. The consequence is that you do not watch movies when you could have without going over the limit.
Type II: H_0 is false, but fail to reject H_0. Claiming that you can afford another movie when your allowance is actually too small. The consequence is that you watch more movies on Netflix and have to pay for going over the allowance.

Score: /4

Problem 5: In an effort to reduce grid lock during rush hour in North Vancouver, city planners have encouraged more people to take public transit and to join car pools. To measure the effectiveness of the campaign, the speeds of 100 motor vehicles were measured during rush hour on Highway 1. The average speed was 13 km/h. Five years ago, before the planners began their campaign, the average speed of vehicles at the same place during rush hour was 11 km/h with a standard deviation of 3 km/h. At the level of $\alpha = 0.05$, can we conclude that rush hour traffic is moving faster than 11 km/h after the city planners' advertising campaign?

a. Define your symbols.

Let $\bar{x} = 13$ km/h be the sample mean speed; μ, the population mean of speed in km/h; $\sigma = 3$ km/h, the standard deviation; $n = 100$, the sample size; H_0, the null hypothesis; H_1, the alternative hypothesis; $\alpha = 0.05$, the level of significance.

b. State the null hypothesis.

H_0: $\mu = 11$ km/h (or H_0: $\mu \leq 11$ km/h)

c. State the alternative hypothesis.

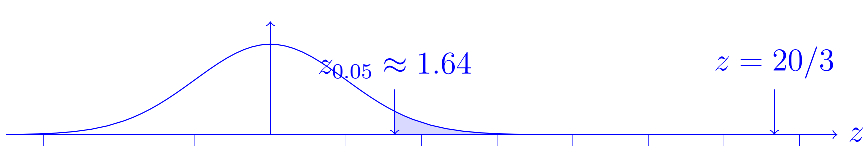
H_1: $\mu > 11$ km/h, also the claim.

d. State the test statistic and compute.

The test statistic with known σ is

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{13 - 11}{3/\sqrt{100}} = \frac{20}{3} = 6\frac{2}{3}$$

e. Draw the critical value on an appropriate graph and locate your test statistic.



f. State your conclusion using everyday language. No technical terms.

There is sufficient evidence (at the significance level of 0.05) to support the claim that rush hour traffic is now moving faster than 11 km/h.

Score: /8

/12

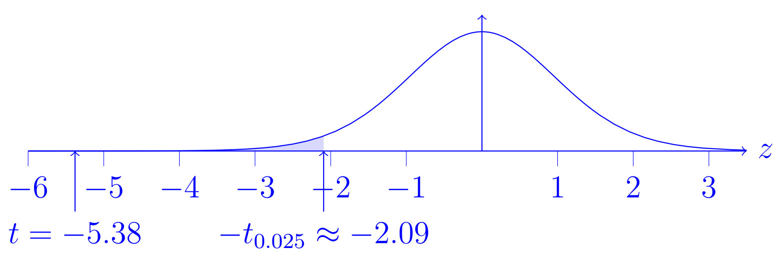
Problem 6: A random sample of 20 international phone calls on a monthly cell phone bill shows the following duration in minutes.

17.5	28.0	38.9	68.4	86.0	99.9	108.6	69.9	16.9	27.3
59.8	10.1	35.3	74.8	87.9	110.7	46.5	18.6	95.3	88.3

Based on this sample, test at the 0.025 significance level, the claim that cell phone users talk an average of less than 100 minutes a month on international calls. Clearly state H_0 and H_1. State the test statistic used and show all steps of hypothesis testing as spelled out by the previous question.

H_0: $\mu \geq 100$ minutes a month.
H_1: $\mu < 100$ minutes a month.
Enter data into a list in the graphing calculator and use $\boxed{\text{1-Var Stat}}$ to calculate sample mean and sample standard deviation.

$$\bar{x} \approx 59.4, \quad s \approx 33.7, \quad n = 20, \quad df = n - 1 = 19, \quad \alpha = 0.025.$$
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{59.4 - 100}{33.7/\sqrt{20}} \approx -5.38$$



Since $-t_{0.025} = \text{invT}(0.025, 19) \approx -2.09$, there is sufficient evidence to support the claim that cell phone users talk less than 100 minutes a month.

Score: /8

/8

Problem 7: A considerable proportion of a person's time is spent working, and satisfaction with the job and satisfaction with life tend to be related. Job satisfaction is typically measured on a four-point scale: very dissatisfied, a little dissatisfied, moderately satisfied, and very satisfied. A numerical scale is created by assigning 1 to very dissatisfied, consecutively up to 4 to very satisfied.

The responses of 226 firefighters and 247 office supervisors yielded the summary statistics:

Firefighter: mean 3.673, sd 0.7235
Office Supervisor: mean 3.547, sd 0.6089

Test at the level of $\alpha = 0.02$ the claim that the mean job satisfaction of firefighters is different from the mean job satisfaction of office supervisors.

Clearly state H_0 and H_1. State the test statistic used and show all steps of hypothesis testing as spelled out by the rush hour traffic question.

Let $n_1 = 226$ be the sample size of the firefighters; $n_2 = 247$ be the sample size of the office supervisors. Let $\overline{x}_1 = 3.673$ be the sample satisfaction mean, $s_1 = 0.7235$ be the sample standard deviation corresponding to the firefighters'. Similarly, let $\overline{x}_2 = 3.547$ be the sample satisfaction mean, $s_2 = 0.6089$ be the sample standard deviation corresponding to the office supervisors'.
The null hypothesis, $H_0 : \mu_1 - \mu_2 = 0$ states that there is no difference between mean job satisfaction of these two groups of people under study.
The alternative hypothesis, $H_1 : \mu_1 - \mu_2 \neq 0$ states that there is a difference between mean job satisfaction. This is the claim of the study.
Since sample data were provided, t-statistic is needed. There is no reason to assume that standard deviation is the same from both populations, so we do not use pooled standard deviation.

$$t = \frac{\overline{x}_1 - \overline{x}_2 - 0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{3.673 - 3.547}{\sqrt{0.7235^2/226 + 0.6809^2/247}} \approx 2.039377$$

Since this is a 2-tailed test, we need two critical values:$\pm\,\mathrm{invT}(0.01, 225) \approx \pm 2.3430$.
The t-statistic does not fall in rejection region, so we fail to reject the null hypothesis.
Thus, we conclude that there is not sufficient evidence to warrant rejection of the claim that mean job satisfaction of firefighters is different from the mean job satisfaction of office supervisors.

Score: /8

/8