1. True or False. Circle the Correct Answer (Highlighted):

T	**F**	a) Suppose (24, 25) is a 95% confidence interval estimate for a population mean μ. Then there is a 95% chance that μ is between 24 and 25.
T	F	b) If 95% confidence intervals are calculated from all possible samples of the give size, μ will be in 95% of these intervals.
T	**F**	c) Two confidence intervals (214.4, 215.6) and (214.1, 215.9) are calculated from the same sample data. The interval (214.4, 215.6) has higher confidence level because it is narrower.
T	**F**	d) When the null hypothesis is not rejected, it is because it is true.
T	**F**	e) If the p-value for a hypothesis test is 0.42, the probability that the null hypothesis is true is 0.42.
T	F	f) Like the normal distribution, the t-distribution is symmetric.
T	F	g)The greater the df, the closer the t-distributions are to the normal distribution.
T	**F**	h) If there is sufficient evidence to reject a null hypothesis at the 5% level, then there is sufficient evidence to reject it at the 1%
T	**F**	i) In conducting a hypothesis test, it is possible to simultaneously make both a Type I error and a Type II error.
T	**F**	j) In medical disease testing with the null hypothesis that the patient is healthy, a Type I error is associated with a false negative; that is the test incorrectly indicates that the patient is disease free.

2. The Indiana State Police wish to estimate the average mph being traveled on the Interstate Highways, which cross the state. If the estimate is to be within ±5 mph of the true mean with 95% confidence and the estimated standard deviation is 25 mph, how large a sample size must be taken?

Solution:

A 95% confidence interval for μ, the true population average speed on the Interstate Highways, is

sample mean ± margin of error, where

The requirement for margin of error is 5 mph . σ=25 mph, and z* = 1.96 for confidence level C=95%.

$$m = z * \frac{\sigma}{\sqrt{n}}$$

5 = 1.96*25/$\sqrt{n}$

n = 96.04 ≈ 97 (round up for all sample size questions)

3. The Indiana Department of Transportation wishes to survey state residents to determine what proportion of the population would like to increase the statewide highway speed limit to 75 mph from 65 mph. How many residents do they need to survey if they want to be at least 99% confident that the sample proportion is within ±0.05 of the true proportion?

Solution:

A 99% confidence interval for p, the true population proportion is

sample proportion ± margin of error, where

$$margin\ of\ error = z * \sqrt{\frac{p(1-p)}{n}}$$

the margin of error is required to be 0.05. z* = 2.575 for confidence level C=99%.

Use p = 0.5 since there is no prior study to estimate it.

0.05 =2.575*$\sqrt{\frac{0.5(1-0.5)}{n}}$

n = 663.06 ≈ 664 (round up for all sample size questions)

4. A laboratory is testing the concentration level in mg/ml for the active ingredient found in a pharmaceutical product. In a random sample of 10 vials of the product, the mean and the standard deviation of the concentrations are 2.58 mg/ml and 0.09 mg/ml. Find a 95% confidence interval for the mean concentration level in mg/ml for the active ingredient found in this product.

Solution:

$\bar{x} = 2.58, s = 0.09$

n = 10 (small sample size ==> use T) ==> DF = 9

t* = 2.262 (use T table, with Confidence Level 95% and DF = 9)

margin of error = 2.262*0.09/$\sqrt{10}$ = 0.644

A 95% confidence interval is (2.58 - 0.064, 2.58+0.064)=(2.516, 2.644)

We are 95% confident that the true average concentration level for the active ingredient found in this product is between 2.516 mg/ml and 2.644 mg/ml

5. The following summary data on proportional stress limits for two different type of woods, Red oak and Douglas fir.

Type of Wood	Sample size	Sample mean	Sample standard deviation
Red oak	50	8.51	1.52
Douglas fir	62	7.69	3.24

a) Find a 90% confidence interval for the difference between true average proportional stress limits for the Red oak and that for the Douglas fir.

Solution:
$\mu1$ = true average proportional stress limits for the Red oak
$\mu2$ = true average proportional stress limits for the Douglas fir
Since both n1 and n2 are larger than 40, we use z* = 1.645 for the 90% confidence interval for $\mu1 - \mu2$.

(8.51 - 7.69) $\pm$1.645*$\sqrt{\frac{1.52\text{^}2}{50} + \frac{3.24\text{^}2}{62}}$ = 0.82$\pm$0.76 = (0.06, 1.58)

b) Interpret your result in part (a).

I am 90% confident that the difference between true average proportional stress limits for the Red oak and that for the Douglas fir is between 0.06 and 1.58.

c) A test of hypotheses is conducted at α=0.10 to determine if the stress limits are the same for the two type of woods. Explain how you can use the confidence interval in part (a) to draw a conclusion in the test of hypotheses.

Solution:

A 90% confidence interval for $\mu 1 - \mu 2$ is (0.06, 1.56) and 0 is outside the interval. This means that 0 is not a plausible value for $\mu 1 - \mu 2$, thus it is not plausible that $\mu 1 = \mu 2$. At significant level α=0.10, we would reject the null hypothesis and conclude that the stress limits are different for the two types of woods at α=0.10 .

6. An article in a medical journal found that percutaneous nephrolithotomy (PN) had a success rate in removing kidney stones of 289 out of 350 patients. The traditional method was 78% effective. Is there evidence that the success rate for PN is greater than the historical success rate? (Use the 7-step procedure to conduct a hypothesis test)

Solution:

Let p denote the population success rate for PN, that is the success rate of PN among all the patients who have gone through this procedure.

Since α is not specified, I will use α = 0.05.

H0: p = 0.78

H1: p > 0.78

Assuming the sample given was a SRS,

N=350 and $\hat{p}$ = 289/350 = 0.8257

N $\hat{p}$ = 289 and N(1- $\hat{p}$) $= 61$ *Both Greater than* 10.

Test statistic:

$$z = \frac{\hat{p} - p0}{\sqrt{\frac{p0*(1-p0)}{N}}} = \frac{0.8257 - 0.78}{\sqrt{\frac{0.78*(1-0.78)}{350}}} = \frac{0.0457}{0.022} = 2.06$$

P-Value = P(Z > 2.06) = 0.0197 < α = 0.05

We reject H0 and conclude that there is evidence that the success rate for RN is greater than the historical success rate.

Note: it is fine if you choose α = 0.01 or α = 0.10 as long as your conclusion is consistent with the significance level α.

7. Fifteen adult males between the ages 35 and 45 participated in a study to evaluate the effect of diet and exercise on blood cholesterol levels. The total cholesterol was measured in each subject initially, and then three months after participating in an aerobic exercise program and switching to a low-fat diet.

The data are shown in the accompanying table. Using Minitab software, Bob and Cathy produced different results for 95% confidence intervals for the mean reduction in the blood cholesterol levels. Answer the questions on the next page.

Table I: Blood Cholesterol Levels for 15 Adult Males

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before	265	240	258	296	251	245	287	314	260	279	283	240	238	225	247
After	229	231	227	240	238	241	234	256	247	239	246	218	219	226	233

Bob's Result:

Two-Sample T-Test and CI: Ex10-41 Before, Ex10-41 After

```
Two-sample T for Before vs After

                 N   Mean  StDev  SE Mean
        Before  15  261.8   25.0      6.4
        After   15  234.9   10.5      2.7


Difference = mu (Before) - mu (After)
Estimate for difference:  26.87
95% CI for difference:  (12.18, 41.55)
T-Test of difference = 0 (vs not =): T-Value = 3.84  P-Value = 0.001  DF = 18
```

Cheryl's Result:

Paired T-Test and CI: Ex10-41 Before, Ex10-41 After

```
Paired T for Before - After

                 N    Mean  StDev  SE Mean
        Before  15  261.80  24.96     6.45
        After   15  234.93  10.48     2.71
Difference      15   26.87  19.04     4.92


95% CI for mean difference: (16.32, 37.41)
T-Test of mean difference = 0 (vs not = 0): T-Value = 5.47  P-Value = 0.000
```

a) What is Bob's 95% confidence interval for the difference of the mean?

(12.18, 41.55)

b) What is Cheryl's 95% confidence interval for the difference of the mean?

(16.32, 37.41)

c) Whose Minitab result is correct? Why?

This is an example of paired design, because the total cholesterol was measured in **each subject** initially, and then three months after participating in an aerobic exercise program and switching to a low-fat diet.

Cheryl's result using Paired T-Test is correct.

d) Using the correct SAS output, carry out a test of hypotheses to determine if the data support the claim that the low-fat diet and aerobic exercise are of value in producing a mean reduction in blood cholesterol levels? Use **α=0.01**.

i) State the null and the alternative hypotheses

μ_d = the true average reduction in blood cholesterol levels in the population.

H0: $\mu_d = 0$

H1: $\mu_d > 0$

ii) Calculate the test statistics based on the correct Minitab result on previous page

test statistics

$$t = \frac{\ddot{d}}{s/\sqrt{n}} = \frac{26.87}{19.04/\sqrt{15}} = \frac{26.87}{4.92} = 5.47$$

iii) Calculate the p-value. Do you reject the null hypothesis?

DF =14 and P-value = 0. Yes, I would reject the null hypothesis.

iv) State your conclusion in terms of the problem.

Yes, data strongly supports the claim that the low-fat diet and aerobic exercise are of value in producing a mean reduction in blood cholesterol levels