**NAME: YOUSAF KHAN**

**ID # 13751**

**TEACHER’S NAME: MS. WAJIHA AMIN**

**SUBJECT: STATISTICAL INFERENCE**

 **FINAL PAPER**

**Q1(a) Suppose a sample of 50 obese patients on a low fat diet lost a mean of 5.5 pounds with a variance of 9 pounds, while another sample of 28 patients on low carb. diet lost a mean of 6.7 pounds with a variance of 16 pounds. Construct 80 % confidence interval for the mean difference of patients on two different diets.**

ANS. 1) n1 = 50 , n2 = 28

 X1=5.5 , X2=6.7

 S1 = 9 **,** S2= 16

 C.I= 80% α=20% = 0.2

 α/2 = 0.1

 tα/2= 1.2928

 S.P = $\sqrt{\left(n-1\right)s1+\left(n2-1\right)s2}/(n1+n2-2)$

 S.P = $\sqrt{\left(49\right)\left(9\right)+\left(27\right)}(16)/76$

 S.P = 3.38

( X1 – X2 ) –tα/2 S.p $\sqrt{\begin{array}{c}\frac{1}{n1}+\frac{1}{n2}\\\end{array}}$ ≤ µ1-µ2 ≤ (5.5-6.7) + (1.2928)

(-1.2) – (1.03) ≤ µ1-µ2 ≤ (-1.2) + (1.03)

 -2.23 ≤ µ1-µ2≤-0.17

**B) Differentiate between z test and t test.**

ANS. 1) The Z test is a calculation that can be used to compare mathematical methods with a sample. T-tests are statistically used to test hypothesis, but they are very useful when we need to determine whether there are significant statistically significant differences between two independent sample groups.

 Usually in stats, you don't know anything about a population, so instead of a Z score you use a T Test with a T Statistic. The major difference between using a Z score and a T statistic is that you have to estimate the population standard deviation.

**Q2:(a) A survey of 250 students indicated that 107 preferred coffee to tea. Determine 90% confidence interval for the proportion of students who preferred coffee.**

ANS. 2(a) n=250 , x=107 . C.I=90% , α = 10% = 0.1 , α/2 = 0.05

1\_0.05= 0.95

 P=x/n = 107/250 = 0.428

P\_Zα/2 $\sqrt{P\left(1-P\right)/n}< π \leq P+Zα/2\sqrt{P(1-P)/n}$

(0.428) – (1.65) . 0.428 (0.572)/250 $<$ π $< $ 0.428+1.65

0.428 – (1.65) (0.0312) $<$ π $<$ 0.428 + (1.65) (0.0312)

0.376 $<$ π $<$ 0.479 ANS

**(b) Briefly discuss point estimate and interval estimate.**

ANS. 2) The point estimation gives us a particular value as an estimate of the population parameter.  Interval estimation gives us a range of values which is likely to contain the population parameter. This interval is called a confidence interval.

**Q3(a)Given  = 1000 , n= 100 sample mean(x) = 870,000, determine the confidence interval for 90% and 98% for mean.**

  = 1000 , n=100 , x=870,000 C.I = 90% C.I=98%

1-0.05 = 0.95

Zα/2 = 1.6 +0.05=1.65

X-Zα/2 . /$\sqrt{n}$ ≤ µ ≤ X + Zα/2 . /$\sqrt{n}$

870,000-1.65 (1000/10) ≤ µ ≤ 870,000 + 1.65 (1000/10)

869835 ≤ µ ≤ 870165

870,000 -2.33 (100) ≤ µ ≤ 870,000 + 2.33 (100)

869,767 ≤ µ ≤ 870,233 answer

 **(b) State what happens to the size of confidence interval as level of confidence increases?**

ANS. 3) Increasing the level of self-confidence increases the risk of error, making the confidence period wider. Lowering self-esteem reduces the risk of error, which in turn reduces confidence.

**Q4:(a)The management of a company is trying to determine annual family medical expenses of its employees. The company wishes to be 95% confident that the mean expenses to be correct within +$50. A previous study indicates the standard deviation of $400. How large the sample size is required for the study?**

Sample size :

N= 1.96 raise to power2 mul by 400 raise to power 2/50rasie to power2

N=245.86

N= 246

**(b) A sample of 12 items has a mean of 7.3 and standard deviation of 2.4.Compute 99% confidence interval for population mean**

**N= 12 ,** x=7.3

= 2.4

C.I= 99% , α= 1%=0.01

α/2=0.005

1-0.005= 0.995

Zα/2=2.58

7.3-2.58 (2.4/3.46) ≤ µ ≤ 7.3 + 2.58 (2.4/3.46)

5.51 ≤ µ ≤ 9.09

**Q 5) FILL IN THE BLANKS:**

1. The probability of any event is between **ZERO** and **ONE**
2. As sample size increases. So the margin of error(e) will **DECREASE**
3. Assuming the variable is normally distributed and the population standard deviation (σ) is unknown, you are conducting a **t-test** for a mean.
4. **. In two to three complete sentences, explain what a confidence interval means (in general), as if you were talking to someone who has not taken statistics.**

ANS 4)

Confidence intervals are a family of statistical techniques that emerge, often in unexpected situations, in your personal and professional lives. Whether you are looking for references to blood tests or risk levels when entering a new business line, confidence intervals enable you to summarize data in a way that reflects the outcome.

1. 2.2 ,0.2 and 20
2. 8, 4 , 81, 80

7 . Estimate expressed by the single value is **POINT ESTIMATE**

8. What is the standard deviation of a sampling distribution called: **(Standard error)**

9. Hypothesis testing and estimation are the two key branches of the field of inferential statistics: **(TRUE)**

10. As a general rule researchers often use **95%** confidence interval