

Solutions

651 Midterm Test (65 minutes)

October, 2009

NAME:

Total number of Marks: /35

There are 6 questions in this paper, do not be deterred, they are all straightforward. Read each question carefully. There are questions on both side of the page. The number of marks for each question are given in brackets. Be smart about how you answer. If you can't answer one question move on the to next and return to the questions you could not do after answering all the other questions!

Rubric: This exam is an open book exam you can use all written materials that you want, normal tables and a calculator.

Write your solutions in the question paper.

GOOD LUCK!

- (1) (i) I draw 300 different samples each sample is of size 50. For each sample I construct a 90% confidence interval (CI) for the mean μ .

On average, how many of the intervals will contain the mean.

$$300 \times 0.9 = 270$$

[1]

- (ii) Suppose you are willing to put 1 out of 200 innocent people in prison. You do the test H_0 : person is innocent against H_A : person is guilty. What is the type I error in this case?

$$\text{Type I error} = \frac{1}{200}$$

[1]

- (iii) Return to question 1(iii), are we able to calculate the type II error in this case?

No

[1]

- (iv) Suppose it is known that the smallest adult is 2 feet tall and the tallest known adult is 8.5 tall. I draw a sample of size 50 people, the average height using this sample is 5.5 feet tall.

Give a 100% CI for the mean adult height.

CI ~~It~~ must for the mean use $[2, 8.5]$

[1]

- (v) Suppose that I draw a random sample of size 40 from a population which has mean μ and variance σ^2 . I evaluate the sample average $\bar{X} = \frac{1}{40} \sum_{i=1}^{40} X_i$. I know that the standard error of the sample average is 0.5. What is the standard deviation of the original population?

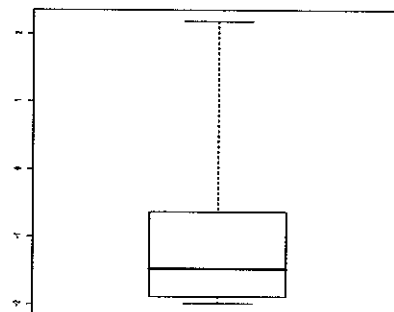
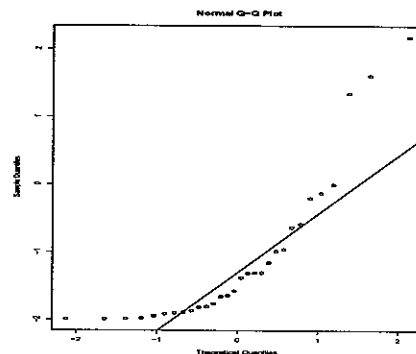
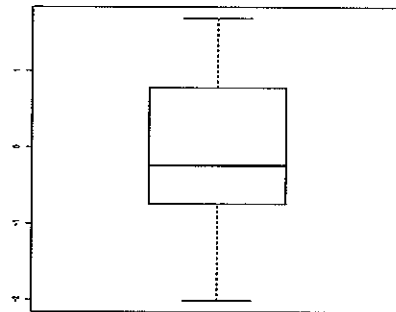
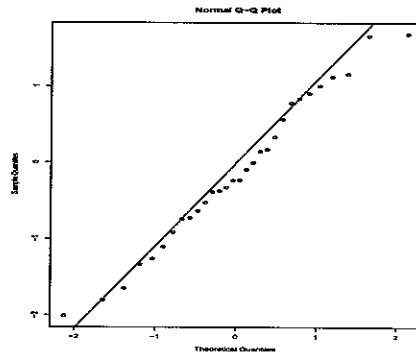
[1]

$$\text{Standard error} = 0.5 = \sqrt{\frac{\sigma^2}{n}} = \frac{s.d}{\sqrt{n}} = \frac{s.d}{\sqrt{40}}$$

$$\Rightarrow s.d = 0.5 \times \sqrt{40} = 3.16$$

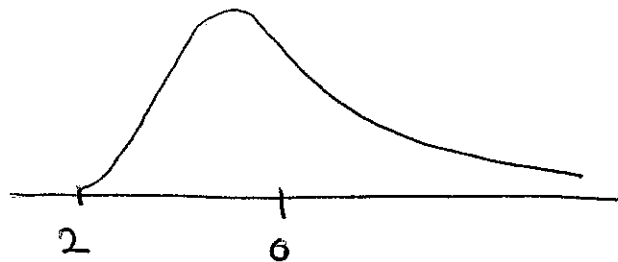
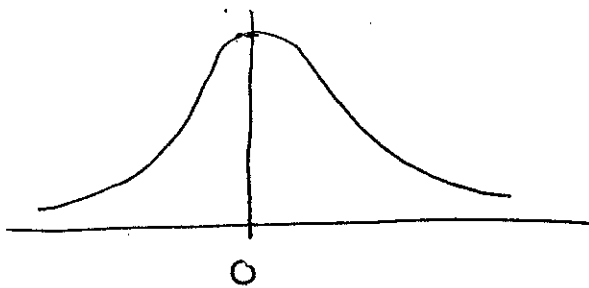
- (2) Two samples each of size 30 are drawn. Sample 1 is drawn from population 1 and Sample 2 is drawn from population 2. A boxplot and QQplot of the samples are given below. The left plots are from sample 1 and the right plot from sample 2.

looks normal



looks right skewed.

- (i) Based on the plots above, sketch the densities (distributions) of population 1 and population 2 (place them side by side). [2]



- (ii) If you were to make:
 (A) 95% confidence intervals for the mean of population 1 based on sample 1.
 (B) 95% confidence intervals for the mean of population 2 based on sample 2.

Which interval do you think is the most accurate 95% CI (in the sense that the sample average is a better approximation of the normal distribution). Give a reason for your answer. [2]

Interval (A) because the original observations are close to normally distributed.

(3) A company claims not to discriminate against women in their employment policy. There are 15 job vacancies and it is observed that one or fewer females get the job (either no women get the job or one women got the job).

- (i) Calculate the probability that one or fewer get the job if the company have a completely fair employment policy (they do not discriminate against men or women) i.e. If 15 randomly selected individuals are employed, the probability that one or less are female. [3]

$$P\{\text{none or female} \mid \text{randomly selected}\} = 0.5^{15}$$

$$P\{\text{one is female} \mid \text{randomly selected}\} = 15 \times 0.5^{15}$$

$$\begin{aligned} P\{S_{15} \leq 1 \mid \text{randomly selected}\} &= 15 \times 0.5^{15} + 0.5^{15} \\ &= 16 \times 0.5^{15} \approx 0 \\ &= 0.00048 \end{aligned}$$

- (ii) Based on the above probability, do you think the company was fair, give a reason for your answer? [2]

Given that the same number of males as females applied for the job and all were equally well qualified, then it appears that company did bias against females since probability is so small.

- (iii) Suppose that the company is a road construction company (a company that builds roads), would you re-evaluate your conclusions in (ii), give a reason for your answer. [1]

The answer to (ii) is based on the assumption that an equal number of females as males applied for the job (and all were equally qualified). In reality, road construction companies tend to attract more male applicants than female applicants. Hence one or less females out of 15 employed could be simply due to their being less female applicants.

(4) (a) Suppose I draw 200 samples (each sample containing 20 observations). For each sample I construct a 90% CI for the mean.

(i) What happens to the size (length) of each interval if I use a 80 observations in each sample rather than 20 observations in each sample? [1]

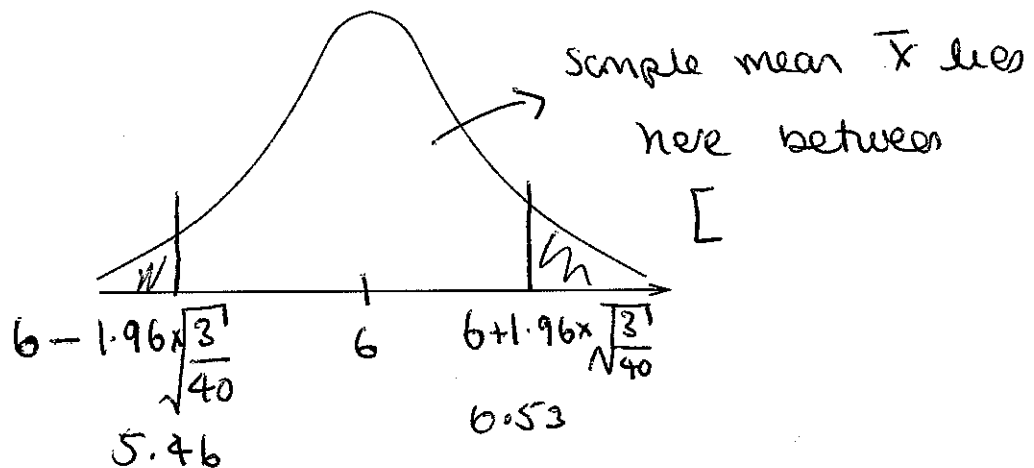
It becomes (50%) ~~also~~ shorter

(ii) What happens to the size (length) of the interval if I construct a 99% confidence interval rather than a 90% confidence interval? [1]

It is longer

(b) Suppose that I want to test the hypothesis H_0 : that $\mu = 6$ against the alternative $H_A : \mu \neq 6$. The population variance is 3 and the sample size 40. I do the test at the 5% level and I am unable to reject the null.

(i) Indicate on the plot below the mean under the null and where the sample mean must lie if we do not reject the null (the interval in which \bar{X} must lie). [2]



(ii) Is this statement correct:

'I could have made a type II error'.

TRUE or FALSE? Give a reason for your answer. [2]

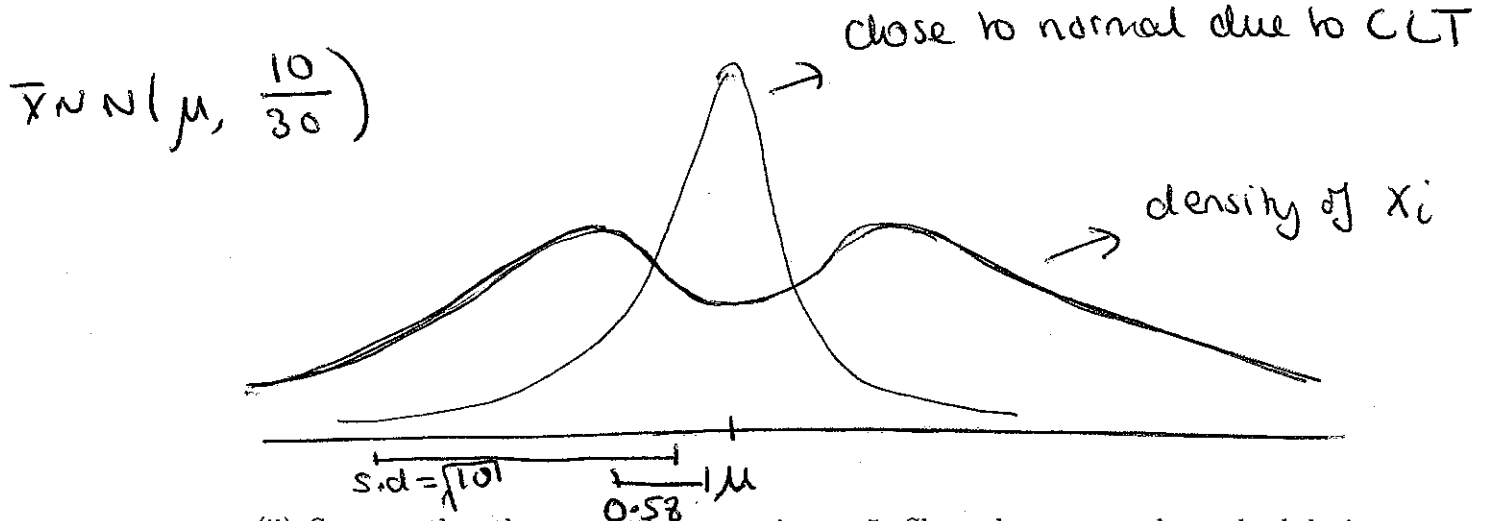
TRUE. ~~False~~

Either I have correctly been unable to reject the null.
Or. I have been incorrectly not been able to reject null
when the alternative is true.

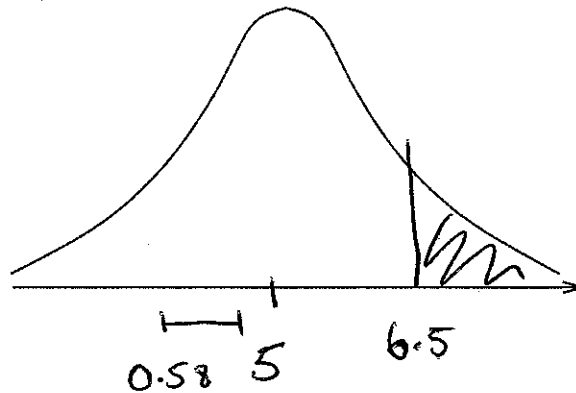
- (5) Suppose that the population mean and variance is μ and 10 respectively. I draw a random sample of size 30 from this population and evaluate the population mean $\bar{X} = \frac{1}{30} \sum_{i=1}^{30} X_i$.

(i) What is the distribution of \bar{X} (give the mean and variance)?

I have made a sketch of the density (distribution) below. Over my sketch make a sketch of the (density) distribution of \bar{X} . [2]



- (ii) Suppose that the population mean is $\mu = 5$. Show the mean and standard deviation of \bar{X} on the plot below. Find the probability that the sample mean \bar{X} is greater than 6.5. [2]



$$P\{\bar{X} > 6.5\} = P\left\{Z > \frac{6.5 - 5}{\sqrt{\frac{10}{30}}}\right\} = P\left\{Z > \frac{1.5}{0.58}\right\}$$

$$= P\{Z > 2.59\} = 0.0047$$

$$\approx \frac{1}{2}\%$$

(5) (continued - from previous page)

(iii) Suppose $\bar{X} = 6.5$. Test the hypothesis that $H_0 : \mu \leq 5$ against $H_A : \mu > 5$ (use $\alpha = 5\%$).

Use calculation from previous page.

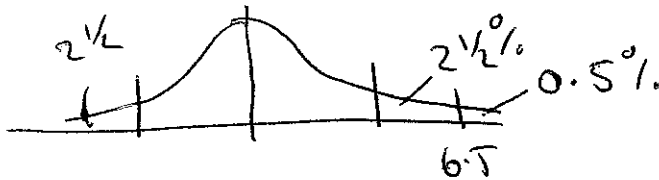
Since \bar{X} is in the alternative and the p-value $\approx 0.005 < 0.05$ we reject the null [1]

(iv) Suppose $\bar{X} = 6.5$. Test the hypothesis that $H_0 : \mu \geq 5$ against $H_A : \mu < 5$ (use $\alpha = 5\%$).

The sample mean $\bar{X} = 6.5$ is not in the alternative.
here there is no evidence to reject the null.

[1]

(v) Suppose $\bar{X} = 6.5$. Test the hypothesis that $H_0 : \mu = 5$ against $H_A : \mu \neq 5$ (use $\alpha = 5\%$).



Since $0.5\% < 2\frac{1}{2}\%$ there is evidence to reject the null.

[1]

- (6) Charlie wants to lose some weight and considers joining 'Dodgy's Dave's Diet club'. But first he wants to make sure that 'Dodgy's Dave's Diet club' claims are correct. He decides to ask a sample of members of 'Dodgy's Dave's diet club' how much weight they have lost. He is interested in the mean weight loss μ (note that if μ is positive there has been weight lost and if μ is negative there has been a weight gain). It is known that the standard deviation of an average person's weight loss is 10 pounds ($\sigma = 10$).

(i) What hypothesis would you advise Charlie to use (state the null and alternative)?

$$H_0: \mu \leq 0 \quad H_A: \mu > 0$$

[1]

- (ii) Suppose that Charlie randomly select 20 members of the club. He decides to do the test at the 5% level. What is the probability that he will reject the null when the average weight loss is greater than 6 pounds ($\mu \geq 6$).

$$P\{\text{reject null} \mid \mu \geq 6\} = 1 - P\left\{Z \leq 1.64 - \frac{10 - 61}{\frac{10}{\sqrt{20}}}\right\} = 1 - P\{Z \leq -1.04\} = 0.85 \quad [2]$$

- (iii) Suppose that Charlie randomly select 20 members of the club, and does the the test at the 5% level. What is the probability he will reject the null, if the mean weight gain is 6 pounds (note there isn't any weight loss!)?

Charlie rejects the null when $\bar{X} \geq 1.64 \times \frac{10}{\sqrt{20}} = 3.67$

$$P\{\bar{X} \geq 3.67 \mid \mu \leq -6\} = P\left\{Z \geq \frac{3.67 - (-6)}{(10/\sqrt{20})}\right\} = P\{Z > 4.32\} = 1 - P\{Z \leq 4.32\} \approx 0$$

Highly unlikely he will reject the null, when there is 6 pound weight gain. [2]

- (iv) Charlie thinks that it is more informative to know the actual mean weight loss. He decides to construct a 95% CI for the population mean μ . He wants the interval to have length 2 pounds. How large a sample size should he choose for the CI to have length 2?

$$2 = 2 \times 1.96 \times \frac{10}{\sqrt{n}} \Rightarrow n = (1.96 \times 10)^2 = 384$$

[2]

$$n \geq 384$$