Solutions

Midterm 2 - STAT 301
Fall 2016

Name:

UIN:

Signature:

Version A:

1. Do not open this test until told to do so.
2. This is a closed book examination, However you may use the cheat sheet provided to you, the t and z tables. You should have no other printed or written material with you on the exam. But scrap paper is allowed.
3. You have 60 minutes to work on this exam. There are 15 multiple choice questions.
4. On the scantron please state the version of exam that you have.
5. You may use a calculator in the exam.
6. If there is no correct answer or if multiple answers are correct, select the **best** answer.
7. Unless stated otherwise, do all tests at the 5% level.
8. If you are unsure of what a question is asking for, do not hesitate to ask the instructor or course assistant for clarification (however we are limited in the amount of help we can offer).
9. Please only give one answer per question (the one that is closest to the solution).
10. Good Luck!!!

(1) Suppose that the height of human males are normally distributed. What proportion of heights will be within 0.5 standard deviations of the mean?

(A) 64% (B) 34% (C) 68% (D) 38%

(E) It is impossible to say without knowing the mean and standard deviation.

look up z-tables

(2) The **maximum** weight a delivery truck can carry is **1000kg**.

$\bar{X}$ normal
mean = 50
$s.e = \frac{10}{\sqrt{18}}$

The weight of an electrical appliance is normally distributed with mean 50kg and standard deviation 10kg. 18 appliances are loaded into the truck, what is the probability that the truck will **exceed** the maximum weight?

$\bar{X} = \frac{X_1 + \dots + X_{18}}{18} > \frac{1000}{18}$

(A) 0.0001% (B) 0.9% (C) 29% (D) 71% (E) 99%

$\frac{\frac{1000}{18} - 50}{(10/\sqrt{18})}$

(3-4) The mean amount of glucose in an expectant female in μ. However, the amount of glucose in a blood sample will vary. The amount of glucose in a blood sample is normally distributed with mean μ and standard deviation σ^2.

A female is diagnosed with gestational diabetes if her mean level $\mu > 140$.

(3) A healthy expectant mother has mean level $\mu = 138$ and the standard deviation $\sigma = 2$. Use the normal distribution to calculate the chance her blood sample will be over 140.

(A) 84.13% (B) 1% (C) 99% (D) 15.86% (E) 2%

(4) 6 blood samples from an expectant mother. Using the data a confidence interval is given below.

One sample T confidence interval:
μ : Mean of population

95% confidence interval results:

Mean	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
μ	142	1.2247449	5	138.85169	145.14831

(A) Since $\bar{X} = 142 > 140$, the data suggests she has gestational diabetes.

(B) Since $138.85 < 140$, the data suggests she does not have gestational diabetes.

(C) Since the interval is both above and below 140, the data does not contain enough information to make a decision.

(D) She clearly has gestational diabetes.

(E) She clearly does not have gestational diabetes.

No Gest. Gest. $\mu > 140$

95% CI = [138.85, 145.14]

(5) In a recent poll a 95% confidence interval for the proportion of the US electrorate who would vote for Hillary Clinton is $[46, 52]\%$. Which statement(s) are correct about the confidence interval?

(A) The margin of error is 6 (B) The margin of error is 3. 1/2

(C) The sample mean is 49 1/2 (D) The sample mean is unknown.

1 (E) (B) and (C).

(6-7) The flight delay times of planes leaving an airport in California are recorded (on time flights or early departures are not included, hence no negative times).

(6) The delay in departures of 6 flights are noted. These delay times are $5.5, 10.5, 13, 22.5, 45, 55$ minutes. The sample mean of this data set is 25.25 and sample standard deviation is 20.2 minutes. Construct a **99%** confidence interval for the mean delay time (use a t-distribution with 5df). CI with negative values makes no sense.

(A) $[25.25 \pm 4.032 \times 20.2]$ (B) $[0, 25.25 + 4.032 \times 8.24]$ (C) $[25.25 \pm 3.65 \times 20.2]$

(D) $[25.25 \pm 3.65 \times 8.24]$ (E) $[0, 20.2 + 4.032 \times 25.25]$.

(7) Based on the plots below, comment on the reliability of the confidence interval constructed in (Q6).

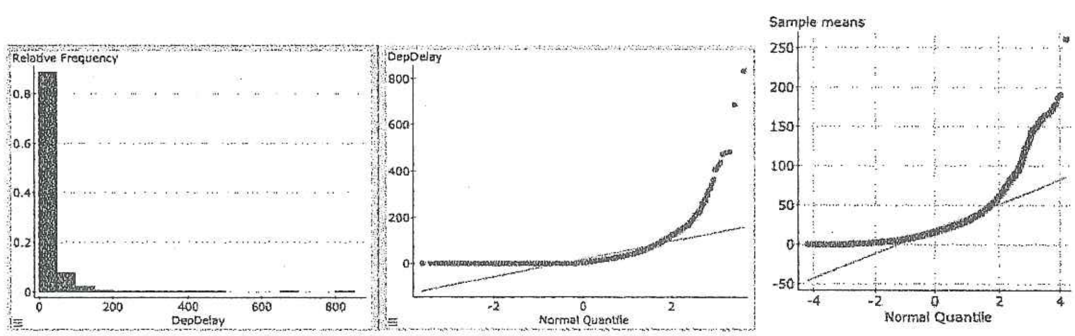


Figure 1: H
istogram, QQplot, QQplot of the sample mean

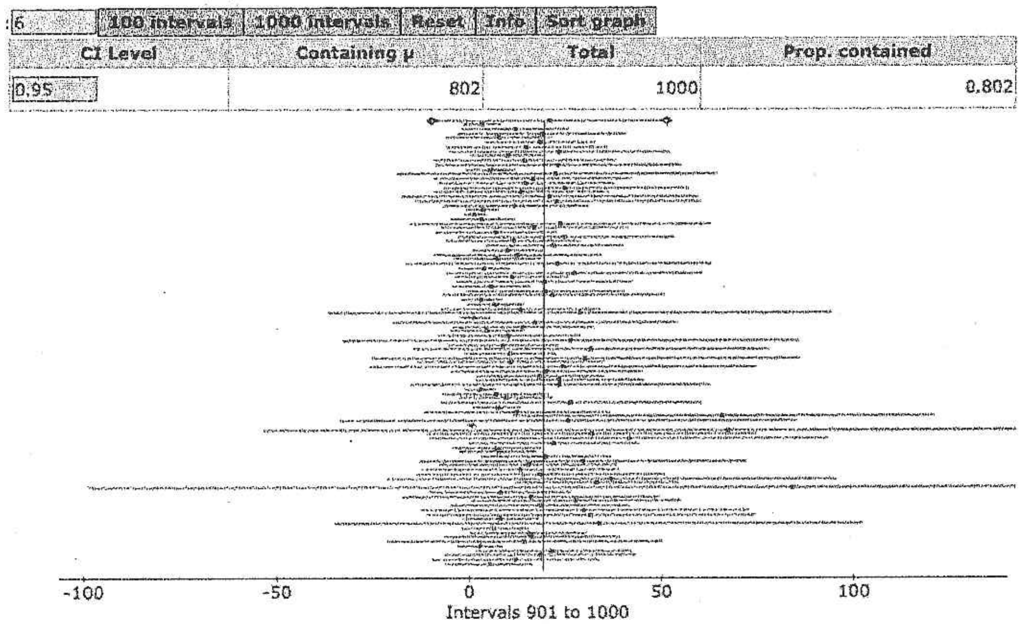
1/2 (A) The delay time data is right skewed.

1/2 (B) The sample mean is right skewed and we do not have 99% confidence in the interval.

(C) The sample mean is normally distributed and we have 99% confidence in the interval.

1 (D) [A] and [B] (E) [A] and [C]

(8) To understand whether a 95% confidence interval constructed from the data (using the t-distribution) is really a 95% confidence interval, 1000 confidence intervals were constructed. The results are summarized in the applet below. Based on the applet, which statement(s) are correct?



No! → (A) The t-distribution is correcting for the lack of normality in the data.

(B) We really do have 95% confidence in this interval.

1 (C) We seem to have 80.2% confidence in this interval.

(D) [A] and [C] (E) [B] and [C].

(9) A health agency wants to construct a 95% confidence interval for the mean height of a 5 year old. The standard deviation of a 5 year old is known to be somewhere between 1 to 3 inches. How many 5 year olds should be sampled to ensure the margin of error is less than 0.5 inches? 1

(A) 16 (B) 44 (C) 62 (D) 139 (E) 156 → 1 (since I did not write minimum sample size).

(10) To understand how the sample mean, sample standard deviation and standard error vary with sample size, random samples of size $n = 3, 10, 50, 200$ were drawn from the population of female students at Texas A&M. A summary of the statistics is given below.

Number	sample mean	sample standard deviation	standard error
3	64.9	1.3	0.752
10	63.9	2.4	0.76
50	63.7	2.47	0.35
200	63.8	2.63	0.186

Just one → simulation

Which statement(s) correctly identify the general behaviour of sample standard deviations and standard errors (you can use the table above to aid your answer).

(A) As the sample size grows the standard error gets smaller. 1/2

(B) As the sample size grows the sample standard deviation grow.

(C) As the sample size grows the sample standard deviation tends to get closer to the population standard deviation. 1/2

(D) (A) and (B) [E] (A) and (C).

1

(11) The mean number of apples eaten in a day (amongst 10 year olds) is known to be 0.4.

Healthy eating advocates want to increase the the number of apples bought. They propose a healthy campaign. After the campaign they sample 50 ten year olds and run the data analysis. The data is summarized below.

One sample T hypothesis test:
μ : Mean of population
$H_0 : \mu = 0.4$
$H_A : \mu > 0.4$

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	0.6	0.084852814	49	2.3570226	0.0112

< 5% reject null
1.12%

From the data what can we conclude about the campaign (use the 5% significance level).

(A) The campaign increased mean apple consumption from 0.4 to 0.6, so it definitely worked.

(B) The sample mean 0.6 is greater than 0.4 however we need to be cautious as the standard error (0.08) is relatively large, thus do not know if the campaign had an impact.

1 (C) The data suggests (at the 5% level) that the campaign increased the mean number of apples consumed.

(D) There is a 1.12% chance the campaign increased apple consumption.

(E) (B) and (D).

(12-13) In soccer, a flip of a coin is used to determine which team attacks which goal (there is one goal on each side of the pitch). The winner of the flip decides which goal it will attack.

One team thinks the coin is unfair and is **more likely to favor heads**. You want to investigate this claim. Let $p =$ the chance of getting a head. The coin is tossed 50 times. The number of heads is 35 out of 50.

(12) What is the hypothesis of interest?

(A) $H_0 : p \geq 0.5$ vs $H_A : p < 0.5$ (B) $H_0 : p \leq 0.5$ vs $H_A : p > 0.5$ 1

(C) $H_0 : p \geq 0.7$ vs $H_A : p < 0.7$ (D) $H_0 : p \leq 0.7$ vs $H_A : p > 0.7$

(E) $H_0 : p = 0.7$ vs $H_A : p \neq 0.7$.

(13) The coin is flipped 50 times. The number of heads was 35 out of 50. The chance of that happening when the coin is fair is 0.3%. What can you conclude about the coin (do all tests at the 5% level)?

$\Rightarrow$ p-value < 5% more heads than tails data suggest coin favors heads.

(A) The sample size of 50 is insufficient to draw any conclusions.

(B) The data suggests that the coin favors heads and $p > 0.5$.

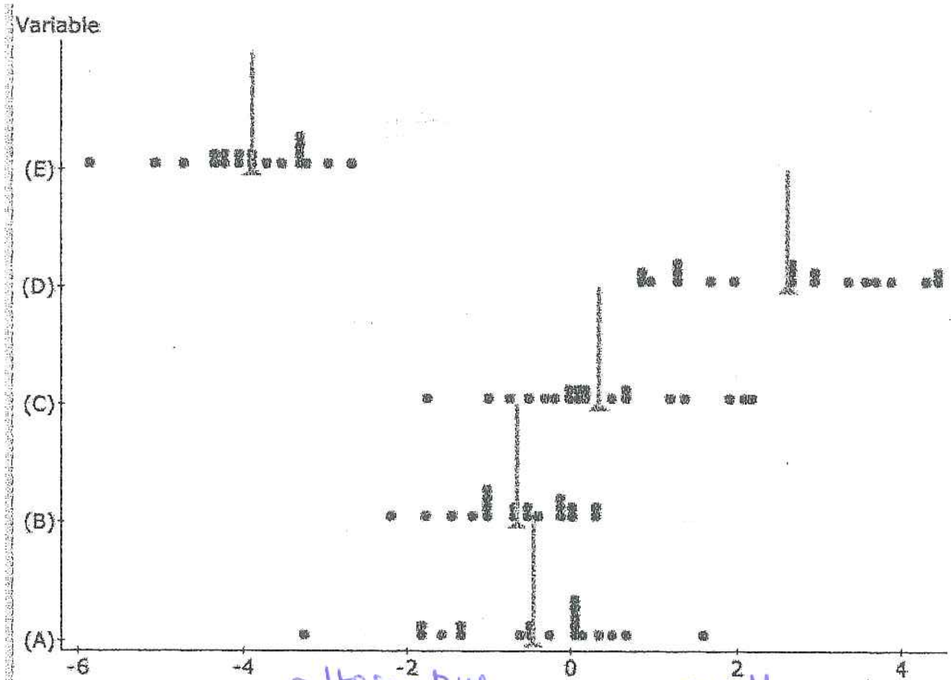
(C) The data suggests that the coin favors tails and $p < 0.5$.

(D) The data suggests that the coin favors tails and $p < 0.6$.

(E) (A) and (C).

Questions 14 and 15 on the next page.

(14-15) 5 samples were drawn each of size 20. The given in the plot below (bottom is labelled (A) going to top which is labelled (E)). Read the hypothesis **carefully**.



alternative null

(14) We test the hypothesis $H_0 : \mu \geq 0$ vs $H_A : \mu < 0$ Which data set has the **smallest** p-value?

(A) (B) (C) (D) (E) → all data in alternative spread very ~~[illegible]~~ small.

(15) We test the hypothesis $H_0 : \mu \geq 0$ vs $H_A : \mu < 0$ Which data set has the **largest** p-value?

(A) (B) (C) (D) (E)

↓ all data in the null.
Could easily obtain such data when the global mean is equal to or ~~[illegible]~~ greater than zero.