Midterm 2 - STAT 301
Fall 2015

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 the cheat sheet, 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!!!

(1-4) The grades in an exam are normally distributed with a (population) mean $\mu = 1100$ and (population/known) standard deviation $\sigma = 200$.

(1) What proportion of students who take this exam will score 1400 **or more**?

(A) 93.3% (B) 98.5% (C) 10% (D) 1.5% (E) 6.68%

(2) A class of **10** student take the exam. What is the distribution of the average (sample mean based on 10) score?

(A) **Right skewed** with mean $\mu = 1100$ and standard error $\frac{200}{\sqrt{10}}$

(B) **Normal** with mean $\mu = 1100$ and standard error $\frac{200}{\sqrt{10}}$

(C) **Normal** with unknown mean and unknown standard error.

(D) **Right skewed** with unknown mean and unknown standard error.

(E) **Normal** with mean $\mu = 1100$ and standard error 200.

(3) What is the chance that the class average will be **1250 or over**?

(A) 0.88% (B) 22.6% (C) 77.3% (D) 99.12% (E) 2.37%.

(4) The chance that the average grade in a class of 10 is **1300 or over** is 0.078%

Astar coaching **randomly samples** 10 students from the population of students who will take the exam. After coaching the average in this class is 1300, using the information at the start of this question, what can we say about Astar coaching?

(A) The probability of this happening by chance is so small (0.078%) it suggests that coaching helped raise the grade.

(B) It is absolutely clear that coaching helped the students.

(C) There is not enough evidence to suggest that coaching helped the students.

(D) There is clear selection bias in the way that Astar selected the group.

(E) [C] and [D].

(5) The 95% confidence interval for the mean based on a sample of size 20 is $[10, 20]$. How large a sample size is required to reduce the margin of error to 1?

(A) We cannot answer without the variance (B) Between 20-100.

(C) 500 (D) 100 (E) 20.

(6-8) The Windchill factor in a certain area is measured over a period of 216 days. The data and the critical values for the t-distribution with 215 degrees of freedom are given below.

Summary statistics:

Column	n	Mean	Variance	Std. dev.	Std. err.
Wind Chill	216	-28.407407	1279.0053	35.763184	2.4333765

probability	0.15	0.10	0.05	0.02	0.01	0.005
t^*	1.03	1.28	1.65	2.06	2.34	2.6

(6) Using the output above, construct a **98%** confidence interval for the mean windchill factor.

(A) $[-28.4 \pm 2.34 \times 2.43]$ (B) $[-28.4 \pm 2.34 \times 35.76]$

(C) $[-28.4 \pm 2.06 \times 2.43]$ (D) $[35.76 \pm 2.34 \times 2.43]$ (E) $[35.76 \pm \times 2.06 \times 28.4]$

(7) Use the output above to test the hypothesis $H_0 : \mu \leq -35$ vs $H_A : \mu > -35$.

(A) The t-value is $t = \frac{-28.407+35}{35.76} = 0.184$. The p-value is less than 0.184% there IS evidence to reject the null at the 5% level.

(B) The t-value is $t = \frac{-28.407+35}{2.44} = 2.7$. The p-value is greater than 99.5% there is NO evidence to reject the null at the 5% level.

(C) The t-value is $t = \frac{-28.407+35}{35.76} = 0.184$. The p-value is greater than 15% there is NO evidence to reject the null at the 5% level.

(D) The t-value is $t = \frac{28.407-35}{2.44} = -2.7$. The p-value is less than 0.5% there is NO evidence to reject the null at the 5% level.

(E) The t-value is $t = \frac{-28.407+35}{2.44} = 2.7$. The p-value is less than 0.5% there IS evidence to reject the null at the 5% level.

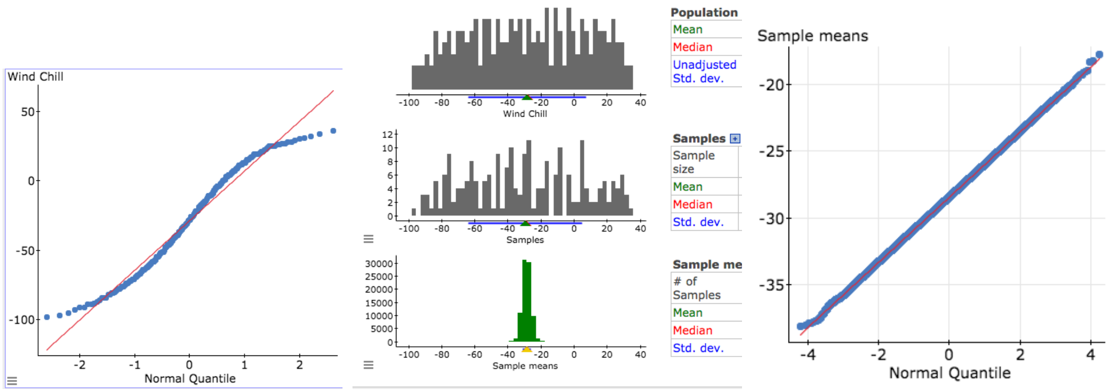


Figure 1: Left: QQplot of Windchill Data. Middle: Distribution of sample mean (sample size $n = 216$). Right: QQplot of the sample mean ($n = 216$).

(8) In order to determine whether we have 98% confidence in the confidence interval constructed in question (6), some plots of the data are given in Figure 1. What can we conclude from these plots?

(A) The Windchill Data is heavy tailed.

(B) The sample mean is normally distributed and we have 98% confidence in the interval.

(C) The sample mean is heavy tailed and we do not have 98% confidence in the interval.

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

(9) In a tomato packing factory, every hour we test the hypothesis $H_0 : \mu = 500g$ vs $H_A : \mu \neq 500g$ to ensure the mean weight of a tomato box does not deviate from 500g (and the machine is functioning correctly). Each test is done at the **10%** level. If the machine is functioning correctly, over 100 hours how many times (on average) will the testing procedure **not** reject the null?

(A) 5 (B) 90 (C) 10 (D) 50 (E) 100.

(10) Suppose in court we test the hypothesis H_0 : innocent vs H_A : guilty.

Which statement(s) are correct?

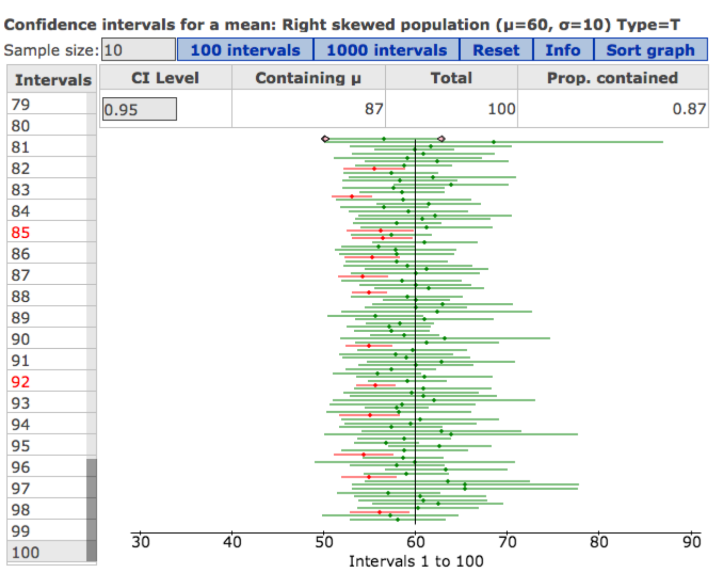
(A) Using the 5% significance level will result in 5% of innocent people being convicted.

(B) Using a 5% significance level will result in 5% of guilty people being convicted.

(C) Using a very low significance level will lead to less innocent people being convicted but also less guilty people being convicted.

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

(11) 100 confidence intervals, at the 95% level (using the t-distribution) for the mean are constructed. The results are summarized in the applet below. Based on the applet, which statement(s) are correct?



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

(B) We seem to have 87% confidence in this interval.

(C) By using the t-distribution we are making the sample mean more normal.

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

(12) In the plot below we have three different data sets. The vertical lines are the means for each data sets (the points are where the data lie). The sample size for each data set is 20. For each data set what is the p-value for $H_0 : \mu \leq 0$ vs $\boldsymbol{H_A : \mu > 0}$.

	Sample1	Sample2	Sample3
A	less than 5%	greater than 50%	greater than 50%
B	greater than 50%	less than 1%	1%
C	less than 5%	less than 1%	greater than 50%
D	greater than 50%	greater than 50%	less than 1%
E	less than 5%	greater than 50%	less than 50%

(13) Hersheys chocolate bar states that that its mean weight is 50g. From previous experience you strongly suspect that the mean weight is LESS than 50g. What is your hypothesis of interest and use the output below to do the test at the 5% level.

Hypothesis test results:
μ : Mean of variable
$H_0 : \mu = 50$
$H_A : \mu \neq 50$

Variable	Sample Mean	Std. Err.	DF	T-Stat	P-value
chocolate	47.839375	0.59971762	15	-3.6027372	0.0026

(A) $H_0 : \mu = 50$ vs $H_A : \mu \neq 50$, the p-value is 0.26% there is evidence to suggest that the mean weight of chocolate bars is **greater** than 50g.

(B) $H_0 : \mu \geq 50$ vs $H_A : \mu < 50$, the p-value is 99.76% there is no evidence to suggest that the mean weight of chocolate bars is **less** than 50g.

(C) $H_0 : \mu \leq 50$ vs $H_A : \mu > 50$, the p-value is 0.13% there is evidence to suggest that the mean weight of chocolate bars is **greater** than 50g.

(D) $H_0 : \mu \geq 50$ vs $H_A : \mu < 50$, the p-value is 0.13% there is evidence to suggest that the mean weight of chocolate bars is **less** than 50g.

(E) $H_0 : \mu \geq 50$ vs $H_A : \mu < 50$, the p-value is 99.87% there is no evidence to suggest that the mean weight of chocolate bars is **less** than 50g.

(14-15) Psychologists want to investigate if people who look dishonest are *given harsher* sentences than people who look honest. The honesty of a person (based only on looks) is rated from 1-8, where 1 is someone who looks very dishonest and 8 is someone who looks very honest (the lower the score the more dishonest the person looks). Let μ_{DR} denote the mean honesty rating of someone on deathrow and μ_{LS} is the mean honesty rating of someone given a life sentence.

(14) What is the hypotheses that the psychologists want to investigate (remember to look how the means are defined)?

(A) $H_0 : \mu_{LS} - \mu_{DR} \leq 0$ vs $H_A : \mu_{LS} - \mu_{DR} > 0$.

(B) $H_0 : \mu_{LS} - \mu_{DR} = 0$ vs $H_A : \mu_{LS} - \mu_{DR} \neq 0$.

(C) $H_0 : \mu_{LS} - \mu_{DR} > 0$ vs $H_A : \mu_{LS} - \mu_{DR} \leq 0$.

(D) $H_0 : \mu_{LS} - \mu_{DR} \geq 0$ vs $H_A : \mu_{LS} - \mu_{DR} < 0$.

(E) $H_0 : \mu_{LS} - \mu_{DR} < 0$ vs $H_A : \mu_{LS} - \mu_{DR} \geq 0$.

(15) The average (sample mean) score for those given a life-sentence is 2.87 whereas the average score for those given the death sentence is 2.76. The statistical analysis showed that the p-value for the test is 0.1%? How to interprete the results of the test at the 5% level?

(A) There is NO evidence to suggest that those with a more dishonest face are given a harsher sentence.

(B) There is evidence to suggest that those with a more dishonest face are given a harsher sentence.

(C) Before using the data to conclude that people who look more dishonest are treated more unfairly, we need to check whether dishonest looking people tend to commit more terrible crimes than more honest looking people.

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