ST 301 (AKI)  Mid2

PLEASE DO NOT OPEN YET!

FIRST NAME:

LAST NAME:

STUDENT ID:

R output (Z-table, t-table)
> round(pnorm(c(-1.644, 0, 1, 1.644, 1.96, 2)), 3)
  [1] 0.050 0.500 0.841 0.950 0.975 0.977

> A=c(0.995,0.99,0.98,0.975,0.95,0.9, 0.85, 0.8)

> round(qnorm(A,mean=0,sd=1), 3)
  [1] 2.576 2.326 2.054 1.960 1.645 1.282 1.036 0.842

> round(qt(A,df=8),3)

> round(qt(A,df=9),3)
  [1] 3.250 2.821 2.398 2.262 1.833 1.383 1.100 0.883

> round(qt(A,df=10),3)
  [1] 3.169 2.764 2.359 2.228 1.812 1.372 1.093 0.879
1. If \( X \) has a normal distribution with \( \mu = -1 \) and \( \sigma^2 = 4 \), find \( b \) such that 
\[ P(X < b) = 0.975 \]
(a) \( b = 1.96 \) (b) \( b = 2.92 \) (c) \( b = 3.92 \) (d) \( b = 6.84 \) (e) None of them

2. The time for an emergency medical squad to arrive at the sports center at the edge of town is distributed as a normal variable with \( \mu = 20 \) minutes and \( \sigma = 4 \). Which arrival period of duration 2 minute is assigned the highest probability by the normal distribution? Which one is correct?
(a) Between 18 and 20 minutes  
(b) Between 19 and 21 minutes  
(a) Between 20 and 22 minutes  
(b) Between 21 and 23 minutes  
(e) None of them

3. A research, interested in the spread of the common cold, believes that 15\% of the population will get a cold during the winter. She will collect data from 30 families of size 4 and count the number of persons \( X \) who get a cold during the winter. Then she consider \( X \) as arising from 120 trials and approximate probabilities be calculated by treating 
\[ \frac{X - np}{\sqrt{np(1-p)}} = \frac{X - 120(.15)}{\sqrt{120(.15)(.85)}} = \frac{X - 12}{\sqrt{15.3}} \]
as a standard normal variable. Which of the following comment is right?
(a) This is right since sample size is large enough.  
(b) This is not right because \( p = .15 \) is near 0.  
(c) This is not right because \( n = 30 \), rather than \( n = 120 \).  
(d) This is not right because results for persons in same family are dependent.  
(e) None of them

4. Which of the following situation is appropriate to apply the normal approximation to the binomial?
(a) \( n=100, p=0.95 \)  
(b) \( n=100, p=0.90 \)  
(c) \( n=100, p=0.80 \)  
(d) All of them  
(e) None of them

5. In a large midwestern university, 40\% of the students live in appartments. If 2400 students are randomly selected, find the probability that the number of them living in apartments will be between 912 and 1008 inclusive.
6. Suppose $Q_1$ and $Q_3$ are the first and third quartiles of the standard normal distribution, respectively. How are these two related?

(a) $|Q_1| = |Q_3|$
(b) $|Q_1| > |Q_3|$
(c) $|Q_1| < |Q_3|$
(d) $Q_3 = 3 \times Q_1$
(e) None of them

Note: This is from p287 [8.6]. Answer is a

7. It is known from past experience that 1% of the tax bills are paid late. If 15000 tax bills are sent out, what is the probability that 165 or more are paid late?

(a) 0.1093  
(b) 0.2186  
(c) 0.8907  
(d) Since 0.01 is too near to 0, the normal approximation to the binomial is inappropriate.
(e) None of them

8. Consider a random sample of size $n = 225$ from a population that has a standard deviation of $\sigma = 60$. What is the probability that the sample mean $\bar{X}$ will differ from population mean $\mu$ by more than 5 units?

(a) 0.2112  
(b) 0.7888  
(c) 0.1056  
(d) 0.8944  
(e) None of them

9. Which of the following is point estimate?

(a) You are standing in line at a theme park and the sign says it takes 10 to 12 minutes to reach the front of the line from here.
(b) A microwave popcorn package reads 'cook for 2 to 3 minutes'.
(c) A report says the average dorm resident consumes 22.5 gallons of water per day.
(d) (a),(b) and (c) are all point estimates.
(e) (a),(b) and (c) are all not point estimates.

10. Suppose you are to verify the claim that $\mu \neq 50$ with a sample of size 100 and you know $\sigma = 5$. If the rejection is $R$: $\bar{X} - 50 \leq -c$ or $\bar{X} - 50 \geq c$, and the level of significance $\alpha=0.02$, then what must numerical value of $c$ be?

(a) c=1.165  
(b) c=1.025  
(c) c=11.65  
(d) c=10.25  
(e) None of them
11. Calculating from a random sample of 64 observations, one obtains the results \( \bar{x} = 80 \) and \( s = 16 \). Determine the \( P \)-value of the test: \( H_0 : \mu = 76 \) against \( H_1 : \mu \neq 76 \).

(a) 0.0228  
(b) 0.0456  
(c) 0.9772  
(d) 0.9544  
(e) None of them

12. Suppose you are to verify the claim that \( \mu \neq 50 \) with a sample of size 100 and you know \( \sigma = 5 \). If you set the rejection to be \( R: \bar{X} - 50 \leq -0.98 \) or \( \bar{X} - 50 \geq 0.98 \), what is the numerical value of type I error \( \alpha \)?

(a) \( \alpha=0.05 \)  
(b) \( \alpha=0.025 \)  
(c) \( \alpha=0.10 \)  
(d) \( \alpha=0.01 \)  
(e) None of them

13. Based on a random sample of size 10 from a normal distribution, an investigator calculates the 95% confidence interval and gets the result (18,19). What is the conclusion of the t-test for \( H_0 : \mu = 18.01 \) against \( H_1 : \mu \neq 18.01 \) at level \( \alpha = 0.05 \)?

(a) \( H_0 \) should be rejected.  
(b) \( H_0 \) should not be rejected.  
(c) \( H_0 \) should be accepted.  
(d) More information are needed.  
(e) None of them

14. If \( X \) is normal distributed with mean \( \mu = 1.5 \) and variance \( \sigma^2 = 4 \). Which one is correct?

(a) \( P[X > 1.6] = 0.4 \)  
(b) \( P[X > 5.5] = P[Z > 1] \), where \( Z \sim N(0,1) \)  
(c) \( \frac{X - 1.5}{4} \) is standard normal.  
(d) \( P[X > 2] = P[X < 1] \)  
(e) None of them

15. Let the number of successes \( X \) have a binomial distribution with \( p = 0.25 \) and \( n = 300 \). What is the approximate probability of \( 67.5 \leq X \leq 75 \)?

(a) 0.34  
(b) 0.6  
(c) 0.5  
(d) 0  
(e) None of them

16. A population has mean 99 and variance 49.

Which one is correct?

(a) \( sd(\bar{X}) = 4.9 \), if sample size is 100.  
(b) \( E(\bar{X}) \) can not be calculated since we don’t know the sample size.  
(c) \( sd(\bar{X}) \) increase with sample size.  
(d) \( sd(\bar{X}) = 0.7 \), if sample size is 100.  
(e) None of them

17. Consider a random sample of size \( n = 100 \) from a population that has a standard deviation of \( \sigma = 20 \). \( \bar{X} \) is the sample mean. Which one is correct?

(a) \( P[\bar{X} - \mu < k] = 0.4 \), then \( k \) should be negative.
(b) \( P[\bar{X} < 3] = 2.7 \)
(c) \( P[1 < \bar{X} - \mu < 2] = 0.6 \)
(d) \( \bar{X} - \mu \) is a standard normal
(e) None of them

18. For estimating a population mean with the sample mean \( \bar{X} \), which one is correct? \( (n = 100, \sigma = 1, \alpha = 0.05) \)
(a) The length of 100(1 - \( \alpha \))% confidence interval is \( z_\alpha \frac{\sigma}{\sqrt{n}} \)
(b) The 100(1 - \( \alpha \))% error margin is 0.196.
(c) The length of 100(1 - \( \alpha \))% Confidence Interval can not be calculated out.
(d) The length of 100(1 - \( \alpha \))% Confidence Interval is just the 100(1 - \( \alpha \))% error margin.
(e) None of them

19. Which one is not a statistic?
(a) Population 10th percentile (b) Sample range (c) Sample standard deviation (d) Sample first quartile (e) Sample median

20. How to calculate \( z_{0.025} \) with R.
(a) \( 2 \times \text{qnorm}(0.5) \) (b) \( -\text{qnorm}(0.025) \)
(c) \( 1/2+\text{qnorm}(0.025) \) (d) \( 1-\text{qnorm}(0.975) \) (e) None of them

Mike observed 100 data \( X_1, ..., X_{100} \) (independent and identically distributed, unknown distribution) where \( E[X_1] = \mu_1 \) and \( Var[X_1] = \sigma^2 \). (large sample case):

Mike did the hypothesis test with \( \alpha = 0.10 \):
\( H_0 : \mu_1 = 200 \)
\( H_1 : \mu_1 < 200. \)

He got \( Z_{\text{obs}} = -1.8 \), and p-value=0.036. (Note: You can trust these numbers. Also \( \bar{x} \) and \( s \) are intentionally omitted).
Nancy observed $Y_1, \ldots, Y_{100}$ where $Y_i = -100 \cdot X_{100-i+1}$, $i = 1, \ldots, 100$. That is, $Y_1 = -100 \cdot X_{100}$, $Y_2 = -100 \cdot X_{99}$, 
$
\ldots$
$Y_{100} = -100 \cdot X_1$.

(For example, if $x_{100} = 3$, $x_{99} = 5$, then $y_1 = -3 \cdot 100$, $y_2 = -5 \cdot 100$.)

Let $E[Y_1] = \mu_2$.

21. Choose the most appropriate answer:
   
   (a) $\mu_2 = \mu_1/100$  
   (b) $\mu_2 = \mu_1/(-100)$  
   (c) $\mu_2 = 100 \cdot \mu_1$  
   (d) $\mu_2 = -100 \cdot \mu_1$  
   (e) None of them

22. Nancy did the hypothesis test with $\alpha = 0.15$:

   $H_0 : \mu_2 = -100 \cdot (200 + 1)$
   $H_1 : \mu_2 < -100 \cdot (200 + 1)$.

   Choose the most appropriate answer:
   
   (a) Her p-value $< 0.036$  
   (b) Her p-value $= 0.036$  
   (c) $0.036 < \text{Her p-value} < 0.964$  
   (d) Her p-value $= 0.964$  
   (e) None of them

23. Nancy did the hypothesis test with $\alpha = 0.15$:

   $H_0 : \mu_2 = -100 \cdot (200 + 1)$
   $H_1 : \mu_2 > -100 \cdot (200 + 1)$.

   Choose the most appropriate answer:
   
   (a) Her p-value $< 0.036$  
   (b) Her p-value $= 0.036$  
   (c) $0.036 < \text{Her p-value} < 0.964$  
   (d) Her p-value $= 0.964$  
   (e) None of them

24. Nancy did the hypothesis test with $\alpha = 0.25$:

   $H_0 : \mu_2 = -100 \cdot (200 + 1)$
   $H_1 : \mu_2 \neq -100 \cdot (200 + 1)$.

   Choose the most appropriate answer:
   
   (a) Nancy’s $Z_{obs} < -1.8$ AND Reject $H_0$  
   (b) Nancy’s $Z_{obs} > 1.8$ AND Reject $H_0$  
   (c) Nancy’s $Z_{obs} < -1.8$ AND Not Reject $H_0$
(d) Nancy’s $Z_{obs} > 1.8$ AND Not Reject $H_0$

(e) None of them

25. Observe $X_1, \ldots, X_{100}$ from $N(\mu = 0, \sigma = 1)$. We estimate $\mu$. We had $\bar{x} = 0.0$, $s = 1.0$. What is 100% C.I. of $\mu$?

(a) (-1, 1)  (b) (-10, 10)  (c) (-100, 100)  (d) $(-\infty, \infty)$  (e) None of them.