Before beginning the exam, please verify that you have 13 pages with 40 questions in your exam booklet. You should also have a decision-tree and formula sheet provided by your TA. Please include your full name and Net-ID on your bubble sheets. Good luck!

1. You want to test whether more than 50% of the UIUC students go to 3 or more bars per weekend. Which of the following tests can you use?
   a. t-test for difference in population means assuming equal variances
   b. t-test for a single population mean
   c. $\chi^2$-test for a single population variance
   d. paired sample t-test for difference in population means
   e. z-test for a single population proportion

2. There is this claim that at least 30% of the students at the U of I belong to some kind of student organization. A survey was conducted and the final result was that 274 students out of 980 interviewed belonged to some student organization. If you perform a z-test to test this claim you would:

\[ z_{0.10} = 1.28 \quad z_{0.05} = 1.645 \]

   a. reject the null hypothesis at 5% of significance, and conclude there is enough evidence to claim that at least 30% of the students belong to some association
   b. reject the null hypothesis at 10% of significance, and conclude there is enough evidence to claim that at least 30% of the students belong to some association
   c. do not reject the null hypothesis at 5% of significance, and conclude there is not enough evidence to claim that at least 30% of the students belong to some association
   d. do not reject the null hypothesis at 10% of significance, and conclude there is enough evidence to claim that at least 30% of the students belong to some association
   e. there is not enough information to answer this question

3. Suppose we have calculated a confidence interval for the difference between two means. Let the population means be denoted by $\mu_1$ and $\mu_2$, the sample means by $\bar{x}_1$ and $\bar{x}_2$, the sample variances by $s_1^2$ and $s_2^2$, and the sample sizes by $n_1$ and $n_2$. What is the midpoint of the estimated interval?
   a. $\bar{x}_1 - \bar{x}_2$
   b. $(\bar{x}_1 / n_1) - (\bar{x}_2 / n_2)$
   c. $\mu_1 / \mu_2$
   d. $\mu_1 - \mu_2$
   e. $(\bar{x}_1 / s_1) - (\bar{x}_2 / s_2)$
Use the following information to answer the next four questions (#4-7)

Let us consider the daily returns of stock A and stock B from Oct, 1992 to Dec, 1996, a total of 1,201 observations. The descriptive statistics for stocks A & B and also sample covariance matrix is given as follows,

<table>
<thead>
<tr>
<th>Stock A: Descriptive Statistics</th>
<th>Stock B: Descriptive Statistics</th>
<th>Stock A</th>
<th>Stock B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean 0.00162</td>
<td>Mean 0.00162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median 0.00162</td>
<td>Median 0.00193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 0.00162</td>
<td>Mode 0.00218</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation 0.0123</td>
<td>Standard Deviation 0.0754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Variance 0.00015</td>
<td>Sample Variance 0.00569</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count 1201</td>
<td>Count 1201</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Sample Covariance Matrix

<table>
<thead>
<tr>
<th>Stock A</th>
<th>Stock B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000151</td>
<td>0.0008</td>
</tr>
<tr>
<td>0.0008</td>
<td>0.005685</td>
</tr>
</tbody>
</table>

4. Assuming that returns of Stock B follow a mound-shaped distribution, what can you conclude about skewness of that distribution?
   a. it positively skewed
   b. it negatively skewed
   c. it is symmetrically distributed
   d. we do not have enough information to infer the skewness of the distribution of stock returns
   e. none of above

5. Which variable would you look at to assess the risk of a stock?
   a. mean
   b. median
   c. mode
   d. standard deviation
   e. none of the above

6. Answer the following question relying only on the Empirical Rule. Based on the mean and standard deviation calculated above for stock A, you can predict that 95% of stock A’s returns will fall between (please calculate up to 3 decimal places)
   a. [-0.004, 0.005]
   b. [-0.035, 0.039]
   c. [-0.023, 0.026]
   d. [-0.011, 0.014]
   e. none of the above
7. What is the sample coefficient of correlation between the returns of stocks A and B?
   a. 0.0008
   b. 0.067
   c. 0.863
   d. 13.231
   e. 54.292

**Use the following information to answer the next two questions (#8-9)**
The dean of CU University claims that students spend on average 2 hours studying daily. In order to check the dean’s claim, a professor conducted a survey with 36 students, where they declared the daily hour spent studying. The mean was found to be 2.5 hours of studying, with standard deviation of 1.5 hours.

\[
\begin{align*}
    z_{0.05} &= 1.64 & t_{0.05, 35} &= 1.690 & t_{0.05, 36} &= 1.688 \\
    z_{0.025} &= 1.96 & t_{0.025, 35} &= 2.030 & t_{0.025, 36} &= 2.028
\end{align*}
\]

8. The test statistic to verify the dean’s claim is equal to
   a. 2
   b. 2.4494
   c. 12
   d. 0.4082
   e. 10

9. At 5% level of significance, the professor would
   a. fail to reject the dean’s claim and conclude there is not enough evidence to say students spend on average 2 hours studying
   b. fail to reject the dean’s claim and conclude there is not enough evidence to say students spend on average less than 2 hours studying
   c. fail to reject the dean’s claim and conclude there is not enough evidence to say students do not spend on average 2 hours studying
   d. fail to reject the dean’s claim and conclude there is not enough evidence to say students spend on average less than 2 hours studying
   e. reject the dean’s claim and conclude there is not enough evidence to say students spend on average less than average 2 hours studying

10. Daily Illini sports columnist has written an editorial comparing the Illini basketball team to the Kansas University basketball team. In general, the lower the opposing team’s field goal percentage, the better the team’s defense is considered to be. The Illini sports columnist wrote that Illini defense, although better, is less consistent than Kansas’ defense. Which test would you perform to test the claim that the Illini defense is less consistent?
    a. t-test for difference in means assuming equal variances
    b. F-test for difference in variances
    c. z-test for difference in means
    d. z-test for difference in proportions
    e. paired sample t-test for difference in means
11. Suppose you want to create a confidence interval for the true mean of the Econ 203 midterm exam scores. Moreover, you don’t want the width of this confidence interval to exceed 2.5. From the previous studies, you know that the population standard deviation is 8.561. What is the minimum sample size you need to create a 99% confidence interval?

\[
z_{0.01} = 2.326 \quad z_{0.005} = 2.576
\]

a. 20  
b. 64  
c. 78  
d. 79  
e. none of the above

12. If the sample covariance between sample X and sample Y is 53.786, what can we say about the relationship between these two samples?

a. they have a strong positive linear relationship  
b. they have a strong negative linear relationship  
c. they have a positive linear relationship  
d. they have a negative linear relationship  
e. we would need to calculate the sample correlation coefficient to be able to determine whether a relationship is positive or negative

13. The 90% confidence interval for the difference in the proportion of women who watch the television show CSI and the proportion of men who watch CSI is \([-0.333, 0.133]\). At the 5% level of significance you can conclude that:

a. the proportion of men who watch CSI is higher than the proportion of women  
b. the proportion of women who watch CSI is higher than the proportion of men  
c. there is a significant difference in the proportion of men and the proportion of women who watch CSI  
d. there is no significant difference in the proportion of men and the proportion of women who watch CSI  
e. there is not sufficient information to make a decision

14. Suppose that, in measuring the variation in the production process of Acme Widgets, you have estimated a 95% confidence interval for the population variance based on a sample of 30 observations. However, in her comments on a draft of your report, your boss says that she thinks your confidence interval is too wide. What actions could you take that would be certain to decrease the width of the interval?

a. use a t distribution instead of a z distribution  
b. decrease the sample size to 25 and lower the confidence level to 90%  
c. increase the sample size to 35 and raise the confidence level to 99%  
d. decrease the sample size to 25 and raise the confidence level to 99%  
e. increase the sample size to 35 and lower the confidence level to 90%
US Global Change Research Center claims that variability of the daily average temperature this winter is different than in the past. To test this claim, Paul collected 10 days’ temperature. The sample of daily average temperatures is the following:

32, 38, 29, 34, 40, 36, 32, 26, 30, 31

Historically, the variance of daily average temperatures was 20.

\[ \chi^2_{0.975, 9} = 2.700 \quad \chi^2_{0.025, 9} = 19.023 \quad z_{0.05} = 1.645 \]

\[ \chi^2_{0.975, 10} = 2.347 \quad \chi^2_{0.025, 10} = 20.483 \quad z_{0.025} = 1.96 \]

15. What is the value of the sample standard deviation?
   a. 4.044
   b. 4.264
   c. 16.36
   d. 18.178
   e. 32.8

16. What is the value of the appropriate test statistic for examining the above claim?
   a. 1.92
   b. 2.13
   c. 8.18
   d. 9.09
   e. 9.90

17. If you were testing this hypothesis at the 5% significance level, what would you conclude?
   a. fail to reject the null hypothesis and conclude there is insufficient evidence to claim that the variability of daily temperature this winter is different than in the past
   b. reject the null hypothesis and conclude that the variability of daily temperature this winter is different than in the past
   c. reject the null hypothesis and conclude there is insufficient evidence to claim that the variability of daily temperature this winter is different than in the past
   d. fail to reject the null hypothesis and conclude the variability of daily temperature in this winter is different than in the past
   e. there is insufficient information to perform the test

18. What is the upper level of a 95% confidence interval for the variability of average daily temperature?
   a. 2.017
   b. 8.600
   c. 50.386
   d. 60.593
   e. 83.469
Use the following information to answer the next two questions (#19-20)

In an agricultural experiment, two expensive high-yield varieties of corn are to be tested and yield improvements are to be measured. The experiment is arranged so that each variety is planted in ten pairs of similar plots. We want to find out if improvement in variety A exceeds the improvement in variety B. We used Excel to create this table using 5% level of significance.

<table>
<thead>
<tr>
<th></th>
<th>Variety A</th>
<th>Variety B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>11.69</td>
<td>10.8</td>
</tr>
<tr>
<td>Variance</td>
<td>7.745</td>
<td>6.369</td>
</tr>
<tr>
<td>Observations</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>0.775</td>
<td></td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>1.565</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.076</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.833</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.262</td>
<td></td>
</tr>
</tbody>
</table>

19. What is the p-value associated with the test you intended to perform?
   a. 1.833
   b. 0.924
   c. 0.152
   d. 0.076
   e. 0.848

20. What is the 95% confidence interval for the true mean of population differences above given that the standard deviation of the differences is $s_D = 1.798$.
   a. [-0.827, 2.047]
   b. [-0.153, 1.932]
   c. [-0.396, 2.176]
   d. [-2.821, 0.047]
   e. there is not enough information available to answer this question
21. Suppose you want to manage your time efficiently and would like to determine whether you work more consistently in the mornings or in the evenings. You start recording times spent working on your stats homeworks and after two months you have enough data to perform an appropriate test. You worked on the stats homework 10 times in the morning (population 1) and 12 times in the evening (population 2). The average time to complete a homework in the morning was 45 minutes with the standard deviation of 17.6 minutes, while the average time to complete a homework in the evening was 45 minutes with the standard deviation of 10.5 minutes. At the 5% level of significance, are you more consistent when working on your homework in the evening?

\[
F_{0.1, 9, 11} = 2.274 \quad F_{0.05, 9, 11} = 2.896 \quad F_{0.025, 9, 11} = 3.588 \\
F_{0.1, 10, 12} = 2.188 \quad F_{0.05, 10, 12} = 2.754 \quad F_{0.025, 10, 12} = 3.374
\]

a. reject the null hypothesis and conclude that you work more consistently in the evening 
b. reject the null hypothesis and conclude that you work more consistently in the morning 
c. fail to reject the null hypothesis and conclude that you work more consistently in the evening 
d. fail to reject the null hypothesis and conclude there is insufficient evidence to claim that you study more consistently in the evening 
e. reject the null hypothesis and conclude there is insufficient evidence to claim that you study more consistently in the evening

22. Which of the following statements is true?
   a. chi-square distribution is symmetric 
b. chi-square distribution is a non-negative distribution 
c. chi-square distribution with 5 degrees of freedom is skewed to right 
d. both b and c 
e. none of the above

23. A Communications professor is curious about knowing whether there is any difference in performance between men and women in her class. She takes a sample of 70 men and 80 women who took her class last semester, and calculates their mean grades (out of 100). She knows from previous years that the standard deviations for men’s and women’s grades are both 12. To test whether there is any significant difference between men’s and women’s grades, which of the following tests is the most appropriate?
   a. z-test for difference in means 
b. t-test for difference in means assuming equal variances 
c. t-test for difference in means assuming unequal variances 
d. paired-sample z-test for difference in means 
e. paired-sample t-test for difference in means
Use the following information to answer the next four questions (#24-27)

Your TA Paul suspects that Canadians spend less money on textbooks than Americans. To test the claim, Paul performs a survey and gathers data from 18 students on the UIUC campus (sample 1) and from 12 students on the University of Toronto campus (sample 2). From the sample, it is found that American students spend on average $360 per semester on books, with a standard deviation of $65, while Canadian students spend an average of $335 per semester on books, with a standard deviation of $55. Assume that the population variances are not equal.

24. To test the claim Paul would perform a hypothesis test. What would the null and alternative hypotheses look like?
   a. $H_0: \mu_1 - \mu_2 \geq 0; H_1: \mu_1 - \mu_2 < 0$
   b. $H_0: \mu_1 - \mu_2 \leq 0; H_1: \mu_1 - \mu_2 > 0$
   c. $H_0: \mu_1 - \mu_2 = 0; H_1: \mu_1 - \mu_2 \neq 0$
   d. $H_0: \sigma_1^2 / \sigma_2^2 = 1; H_1: \sigma_1^2 / \sigma_2^2 \neq 1$
   e. $H_0: \sigma_1^2 - \sigma_2^2 \geq 0; H_1: \sigma_1^2 - \sigma_2^2 < 0$

25. What is the value of the point estimate?
   a. 1.133
   b. 1.075
   c. 20
   d. 25
   e. 45

26. Ignoring previous calculations suppose that the correct test statistic was 2.377. What is the conclusion about Paul’s claim at the 5% significance level?

   \[
   t_{0.05, 26} = 2.060 \quad t_{0.025, 26} = 2.385 \quad t_{0.01, 26} = 2.787 \\
   t_{0.05, 28} = 2.052 \quad t_{0.025, 28} = 2.373 \quad t_{0.01, 28} = 2.771
   \]

   a. reject the null hypothesis and conclude there is insufficient evidence to claim that Canadian students spend less on textbooks than American students
   b. fail to reject the null hypothesis and conclude there is insufficient evidence to claim that Canadian students spend less on textbooks than American students
   c. fail to reject the null hypothesis and conclude that Canadian students spend less on textbooks than American students
   d. fail to reject the null hypothesis and conclude that spending on books is different among American and Canadian students
   e. reject the null hypothesis and conclude that Canadian students spend less on textbooks than American students
27. Suppose that you, on the other hand, wanted to test whether the American students spend more than $25 on average on books than Canadian students. What would be the value of the test statistic to perform such a test?
   a. -8.733
   b. -1.133
   c. 0
   d. 1.133
   e. 8.733

Use the following information to answer the next three questions (#28-30)
A bus company, specialized in transporting students from Chambana to Chicago on weekends, wants to extend its business and acquire new buses to serve the students. The manager believes that this expansion would be profitable only if it is found that at least 35% of the students travel to Chicago on a regular basis. After conducting a survey with 120 students, 45 of them answered that they go to Chicago area regularly.

\[ z_{0.05} = 1.645 \quad z_{0.025} = 1.96 \]

28. Which of the following should the manager consider as null and alternative hypotheses?
   a. \( H_0: p = 0.35; H_1: p < 0.35 \)
   b. \( H_0: p \neq 0.35; H_1: p < 0.35 \)
   c. \( H_0: p \neq 0.35; H_1: p > 0.35 \)
   d. \( H_0: p = 0.35; H_1: p > 0.35 \)
   e. \( H_0: p = 0.35; H_1: p \neq 0.35 \)

29. What is the value of the appropriate test statistic?
   a. 0.5656
   b. 0.5742
   c. -0.5656
   d. 0.5000
   e. 0.3872

30. Ignoring the answer that you got in the previous question, suppose that the correct test statistic was 1.074. What is your conclusion at the 5% level of significance?
   a. the manager rejects the null hypothesis and should acquire new buses
   b. the manager rejects the null hypothesis and should not acquire new buses
   c. the manager fails to reject the null hypothesis and should acquire new buses
   d. the manager fails to reject the null hypothesis and should not acquire new buses
   e. since we need excel to compute the p-value, it is impossible to conclude any of the above
31. Suppose that you are performing the following test: $H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1; \ H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$ and you get the test statistic $F = 0.21$. Which shaded area (designated by the arrow) is the appropriate p-value for this test?

*Note: the left end-point of all the distributions is zero.*
Use the following information to answer the next three questions (#32-34)

You are an owner of a bookstore at the University of Michigan campus, but you are debating whether to transfer the bookstore to the U of I campus. You will transfer the store to the U of I campus only if the average spending on books in Urbana-Champaign is greater than at the University of Michigan. An earlier survey concluded that the average per capita spending on books at the University of Michigan is $200 per year. A sample of 64 students was interviewed in Urbana-Champaign and the average spending on books was $212, with a standard deviation of $72.

\[
t_{0.10, 63} = 1.29 \quad t_{0.05, 63} = 1.67 \quad z_{0.10} = 1.28 \quad z_{0.05} = 1.64
\]

32. What are the null and the alternative hypotheses in this case, and which test should you perform?
   a. \( H_0: \mu = 212 \) and \( H_1: \mu \neq 212 \); and a z-test
   b. \( H_0: \mu \leq 212 \) and \( H_1: \mu > 212 \); and a t-test
   c. \( H_0: \mu \geq 200 \) and \( H_1: \mu < 200 \); and a t-test
   d. \( H_0: \mu \leq 200 \) and \( H_1: \mu > 200 \); and a t-test
   e. \( H_0: \mu = 200 \) and \( H_1: \mu \neq 200 \); and a t-test

33. What is the value of the appropriate test statistic?
   a. -1.333
   b. 0.019
   c. 1.333
   d. 10.67
   e. 23.556

34. Ignoring the result of the previous question, assume that the correct test statistic was 1.999. What would your decision be at the 10% level of significance?
   a. reject the null hypothesis and conclude that the average per capita spending on books is greater at the U of I campus than at the University of Michigan campus, so open the store
   b. fail to reject the null hypothesis and conclude there is not enough evidence to claim that the average per capita spending on books is greater at the U of I campus than at the University of Michigan campus, so do not open the store
   c. fail to reject the null hypothesis and conclude that the average per capita spending on books is greater at the U of I campus than at the University of Michigan campus, so open the store
   d. reject the null hypothesis and conclude there is not enough evidence to claim that the average per capita spending on books is greater at the U of I campus than at the University of Michigan campus, so open the store
   e. reject the null hypothesis and conclude that the average per capita spending on books is greater at the U of I campus than at the University of Michigan campus, so do not open the store
Use the following information to answer the next three questions (#35-37)

An economic journal announces that finding jobs for University of Chicago grads was harder than for New York University grads in 2003. Two independent samples of 800 graduates from each university were selected. In the Chicago sample (sample 1), 40 graduates were unemployed, while in the New York sample (sample 2), 32 graduates were unemployed.

35. The test statistic for testing the economic journal’s claim will have the following distribution:
   a. t distribution with 1598 degrees of freedom
   b. z distribution
   c. t distribution with 32 degrees of freedom
   d. F distribution with 32 degrees of freedom
   e. chi-squared distribution with 1598 degrees of freedom

36. What is the value of the test statistic for this test?
   a. -0.918
   b. -0.842
   c. 0.842
   d. 0.965
   e. 2.01

37. Ignoring the result you got in the previous question, suppose that the correct test statistic was 1.173. What is your conclusion at the 10% level of significance?
   a. reject the null hypothesis and conclude there is not enough evidence to support the journal’s claim
   b. reject the null hypothesis and conclude that the journal’s claim is correct
   c. fail to reject the null and conclude that the journal’s claim is correct
   d. fail to reject the null hypothesis and conclude there is not enough evidence to support the journal’s claim
   e. accept the null hypothesis, and conclude that the journal’s claim is correct
38. Paul and Tony meet after the soccer match and start discussing their favorite teams. Paul claims that Manchester United FC is a better team than FC Real Madrid since Manchester United has more season-ticket holders. Tony, on the other hand, claims that Real Madrid is a better team since the occupancy rate at the stadium is higher at Real Madrid home games than at the Manchester United home games. If you wanted to test Paul’s claim, which test would be the most appropriate given that you are limited to the choice of test to the following?
   a. t-test for difference in means assuming unequal variances
   b. z-test for difference in proportions
   c. z-test for difference in means
   d. paired sample t-test for difference in means
   e. F-test for difference in variances

**Use the following information to answer the next two questions (#39-40)**

Two students are arguing over who has better pizza, Papa John’s or Domino’s. In particular, they both think their preferred pizza chain has bigger “Large” pizzas. They buy ten of each pizza, and calculate the following: Papa John’s pizzas (sample 1) had a mean diameter of 14.1” with a standard deviation of 0.2”, while Domino’s pizzas (sample 2) had a mean diameter of 14.2” with a standard deviation of 0.3”. Moreover, from previous studies the students know that variances of the two populations are equal.

39. What can you say about the test statistic?
   a. it equals -0.877 and follows a t distribution with 18 degrees of freedom
   b. it equals -0.877 and follows a t distribution with 19 degrees of freedom
   c. it equals -0.447 and follows a t distribution with 9 degrees of freedom
   d. it equals -0.1 and follows a t distribution with 18 degrees of freedom
   e. it equals -0.1 and follows a t distribution with 19 degrees of freedom

40. Now suppose that you are skeptical about the equal population variances assumption used throughout this set of questions. So, you perform the appropriate test. What is your conclusion at the 10% level of significance?

   \[
   F_{0.1, 9, 9} = 2.440 \quad F_{0.05, 9, 9} = 3.179 \quad F_{0.025, 9, 9} = 4.026 \\
   F_{0.1, 10, 10} = 2.327 \quad F_{0.05, 10, 10} = 2.978 \quad F_{0.025, 10, 10} = 3.717
   \]

   a. there is insufficient evidence to conclude that the population variances are equal, so the above assumption of equal variances is correct
   b. there is insufficient evidence to conclude that the population variances are equal, so I would have to re-evaluate the above answers
   c. there is enough evidence to conclude that the population variances are unequal, so the above assumption of equal variances is correct
   d. there is enough evidence to conclude that the population variances are unequal, so I would have to re-evaluate the above answers
   e. there is insufficient evidence to conclude that the population variances are unequal, so the above assumption of equal variances is correct
FORM B

1. E
2. C
3. A
4. B
5. D
6. C
7. C
8. A
9. C
10. B
11. C
12. C
13. D
14. E
15. B
16. C
17. A
18. D
19. D
20. C
21. D
22. D
23. A
24. B
25. D
26. E
27. C
28. D
29. B
30. D
31. D
32. D
33. C
34. A
35. B
36. D
37. D
38. A
39. A
40. E