

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.

Directions

The final exam will consist of several application-type questions related to the following topics we've covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The dataset that you will examine is introduced below including actual questions that will be on the exam. You should load these data into R (from the class website) and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

- The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
- The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write *any other* material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam (29% of your overall grade)!!

Data Set – Loon1

Dr. Jim Paruk has studied the morphology of common loons (*Gavia immer*), primarily their bills, collected from a wide variety of locations in North America. Specifics of the variables recorded are documented in the **Loon1.txt** data file (open the data file outside of R to read comments at the top). *You should ignore the State variable in all of your analyses.* You should prepare results for each of the following items ...

1. Univariate EDA for loon weight, bill length, and region.
2. Bivariate EDA for tarsus length and weight, bill length and weight
3. Linear regression results (equation results and r^2) for predicting weight from bill length.
4. Results for testing the following research hypotheses (use 5% level for each)
 - a. The mean weight is different between **KNOWN** male and female loons.
 - b. The distribution of individuals into the three regions differs between **KNOWN** male and female loons.
 - c. The mean culmen length for male loons is greater than 66 mm.

Directions:

You may have a pencil, a calculator, and your R output document for the **Loon1.txt** file on your desk. All other materials should be fully stored out of sight and your computer should be turned off.

Do not write anything on your document of results except to add labels – e.g., “Output #1” or “Figure 1” – for referring to when answering the questions below. When you are finished with the exam, you should staple this sheet, your handwritten answers, and your document of R results together, in that order.

You should answer all questions below with as much information as necessary to fully answer the question. All answers should be completed by using and referring to specific R output. Some questions require further calculations for which you are allowed to use your calculator. You are not allowed to make any further calculations in R. Your answers should be legibly handwritten on the sheets of paper provided, clearly labeled with the question number, and, *when marked by an asterisk*, written with complete sentences.

11 Steps for any Significance Test

1. [1] state the rejection criterion (α),
2. [2] state the null and alternative hypotheses to be tested – define the parameter,
3. [1] determine which hypothesis test to use – thoroughly explain why,
4. [1] collect the data (address type of study and randomization),
5. [2] check all necessary assumptions – explain how you tested the validity,
6. [1] calculate the appropriate statistic(s),
7. [2] calculate the appropriate test statistic,
8. [2] calculate the p-value,
9. [1] state rejection decision,
10. [2]* summarize your findings in terms of the problem, and
11. [2]* If reject H_0 , compute a **100(1- α)% confidence region** for the parameter.

Questions:

1. [3pts] Identify what type of variable each of the following is: weight, bill.len, and region.
2. [5pts]* Perform a thorough EDA for loon weight.
3. [2pts]* Perform a thorough EDA for the region variable.
4. [5pts]* Perform a thorough EDA for the relationship between tarsus length and weight.
5. [2pts]* Interpret the slope of the linear regression that you performed.
6. [2pts] Predict the weight of a loon if the bill length equals the median bill length.
7. [2pts] What proportion of the total variability in weight is explained by knowing the bill length?
8. [15pts] Test that the distribution of individuals into the three regions differs between **KNOWN** male and female loons.
9. [15 or 17 pts] Test that the mean weight is different between **KNOWN** male and female loons.
10. [8 pts]* Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important.