1. (23pt) The weights of the 7 kittens are displayed in the table below. It is assumed that the weights of the kittens is normally distributed.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Kitten number | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Weight (lbs) | 12.1 | 4.3 | 6.4 | 12.7 | 15.3 | 5.6 | 16.2 |

(=72.6, =142.0743)

1. (3pt) Compute the sample mean and the sample variance of the kitten’s weight.
2. (4pt) Compute a 95% confidence interval for if it is assumed that the standard deviation of the population is
3. (4pt) Compute a 95% confidence interval for if it is assumed that the standard deviation is unknown.
4. (4pt) Perform the hypothesis test for the following hypotheses : vs . Assume that the standard deviation of the population is and use α =0.05.
5. (4pt) Perform the hypothesis test for the following hypotheses : vs . Assume that the standard deviation is unknown and use α =0.05.
6. (4pt) When is known as in (d), if a level 0.05 test is used, what is , the probability of a type II error when ? (Represent in the form of )
7. (6pt) Eight properties were assessed by two tax assessors. The assessments (in thousands of dollars) are shown in the table.

|  |  |
| --- | --- |
| Assessor A | 76.3 88.4 80.2 94.7 68.7 82.8 76.1 79.0 |
| Assessor B | 75.1 86.8 77.3 90.6 69.1 81.0 75.3 79.1 |

Based on the following R output, answer the questions. (You do not need to compute from the data, only use following R output.)

1. (2pt) Fill in the (1)-(3) of the following R statement.
2. (3pt) Do the data provide sufficient evidence to indicate that assessor A tends to give higher assessments than assessor B when ? State the null hypothesis and alternative hypothesis, p-value and the conclusion.
3. (1pt) Construct the 95% confidence interval of the difference of the means of Assessor A and Assessor B.

> a <- c(76.3, 88.4, 80.2, 94.7, 68.7, 82.8, 76.1, 79.0)

> b <- c(75.1, 86.8, 77.3, 90.6, 69.1, 81.0, 75.3, 79.1)

> \_\_\_(1)\_\_\_(\_\_\_(2)\_\_, \_\_\_(3)\_\_)

Welch Two Sample t-test

t = 0.39963, df = 13.679, p-value = 0.6956

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-6.513393 9.488393

sample estimates:

mean of x mean of y

80.7750 79.2875

1. (10pt) In order to investigate the relationship between the Consumer Price Index(CPI) and the Dow Jones Industrial Average(DJA), a random sample of ten months was selected from several years in the 1970s. The following shows the corresponding DJA and CPI data.

|  |  |  |
| --- | --- | --- |
| DJA | x | 660 638 639 597 702 650 579 570 725 738 |
| CPI | y | 13.0 14.2 13.7 15.1 12.6 13.8 15.7 16.0 11.3 10.4 |

( =30687.6, =29.716,

=-941.04)

1. (5pt) Fit a least-squares line to the data. (x is a predictor variable and y is the response variable)
2. (3pt) Calculate the correlation coefficient r.
3. (2pt) Predict the CPI when DJA is 690.
4. (17 pt) Suppose we have the following data set. x is the predictor variable and y is the response variable.

|  |  |
| --- | --- |
| x | 1 -5 12 5 7 0 -6 2 |
| y | 2 -10 18 9 12 -1 -12 2 |

The linear regression analysis was performed by R as follows. Answer the questions based on the R output.

(You do not need to compute from the raw data, only use following R output.)

1. (2 pt) Fill in (1)-(3) of the following R statement for the linear regression model.
2. (2 pt) Fill in (4) of the following R output.
3. (2 pt) What is the regression line?
4. (2 pt) Predict y when x=1.
5. (2 pt) What is the proportion of the variance in y explained by regression?
6. (4 pt) Construct a 95% confidence interval for the slope of the regression line,
7. (3 pt) We want to perform the hypothesis test with α =0.05: :=0 vs :0. What is the test statistic value, p-value, and conclusion. You may answer based on the following R output.

> x <- c(1, -5, 12, 5, 7, 0, -6, 2)

> y <- c(2, -10, 18, 9, 12, -1, -12, 2)

> reg <- \_\_(1)\_\_(\_\_(2)\_\_ ~ \_\_(3)\_\_)

> summary(reg)

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.93651 0.43329 \_\_(4)\_\_ 0.0739 .

x 1.71825 0.07272 23.628 3.77e-07 \*\*\*---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.154 on 6 degrees of freedom

Multiple R-squared: 0.9894, Adjusted R-squared: 0.9876

F-statistic: 558.3 on 1 and 6 DF, p-value: 3.772e-07

1. (11 pt) Suppose we want to examine the safety of compact cars, midsize cars, and full-size cars. We collected a sample of three for each of the treatments (cars types). Using the data provided below, test whether the mean pressure applied to the driver’s head during a crash test is equal for each types of car. Use α =0.05.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Compact cars  Midsize cars  Full-size cars | 643, 655, 702  469, 427, 525  484, 456, 402 | 2000  1421  1342 | 666.67  473.67  447.33 |
|  |  | 4763 | =529.22 |

1. (2 pt) State the null and alternative hypothesis
2. (4 pt) Compute the sum of squares due to type of cars.
3. (3 pt) Fill in the blank space of the ANOVA table. If you could not compute the sum of squares due to type of cars in b), assume that it is 80000.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source of  Variation | df | Sum of  Squares | Mean Square | F |
| Type  Error  Total | \_\_(1)\_\_  \_\_(2)\_\_  8 | \_\_\_(3)\_\_  \_\_\_(4)\_\_  96304 | \_\_\_(5)\_\_  \_\_\_(6)\_\_ | \_\_\_(7)\_ |

1. (2 pt) Perform the hypothesis test based on the above table and the critical value of the F distribution.
2. (10 pt) The side effects of a new drug are being tested against a placebo. A simple random sample of 565 patients yields the results below.

|  |  |  |  |
| --- | --- | --- | --- |
| Result | Drug | Placebo | Total |
| Coughing | 36 | 13 | 49 |
| No coughing | 254 | 262 | 516 |
| Total | 290 | 275 | 565 |

1. (4pt) Calculate the expected number of patients in each cell of the above table when coughing is assumed to be independent of treatment.
2. (3 pt) Calculate test statistic.
3. (3 pt) At a significance level of =0.05, is there enough evidence to conclude that the treatment is independent of the side effect of coughing?
4. (16 pt) Consider a random sample collected from a continuous distribution with density

for 0<x<1

1. (6 pt) Use the method of moments to obtain an estimator of
2. (6 pt) Obtain the maximum likelihood estimator of .
3. (4 pt) Suppose that we have 3 observations, , , Estimate by the method of moments and by the method of maximum likelihood

Cf)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| z | 0.5 | 1 | 1.28 | 1.645 | 2 | 2.33 | 3 |
| ) | 0.6915 | 0.8413 | 0.9 | 0.95 | 0.9772 | 0.99 | 0.9984 |

, , ,

, ,

, , ,

, , ,

, , ,

=9.49, =5.99, =3.841, =11.143, =7.378, =5.024