

[Home](#) / [My courses](#) / [Capstone Project | Omirgaliev Ruslan](#) / [Week 10](#) / [Final FX](#)

Time left 1:48:40

Question **1**

Not yet answered

Marked out of 12.00

Suppose that you have the following results of survey. Now it is necessary to analyze obtained survey by using Simple Random Sampling and Stratified Random Sampling. Recommended tool for using is MS Excel.

(Ignore the fpc and the clustering in calculating the standard error.)

P.S. when you are going to write your answer into answer sheet, please round up to 2 digits after floating point

1) Assume that you're going to do Simple Random Sampling (SRS) for above dataset.

Compute a mean:

P.S here you need to drag-and-drop digit-by-digit. For example, if your answer is 35.13 then you need to drag 3, then 5, then . and then 1 and 3

2) Compute a standard error for SRS:

3) Now compute 95% of confidence interval. Please note that t-value in this case is equal to 2.04

Upper limit for SRS:

Lower limit for SRS:

P.S You need to take **rounded** answers from 1st and 2nd questions

Now for Stratified part

1) What is the value of  $W_h$

2) Compute a mean:

3) Compute a standard error for Stratified part:

4) Compute d-value:

Hint: d- value is a ratio of standard error for statified over standard error for SRS

5) Compute d-squared:

6) Compute  $N_{eff}$

P.S you need to take exact anwer from 5-th step rounded up to 2 digits after floating point.

1234567890,

Question **2**

Not yet answered

Marked out of 10.00

Suppose that you're going to run linear regression with some input features and 1 output feature. Your hypothesis is linear

```
X = np.c_[np.ones(df.shape[0]), df[['X1', 'X2', 'X1^2', 'X1^3', 'X2^2', 'X2^3']].values]
Y = df['Y'].values.reshape(-1, 1)
```

Firstly it is necessary to normalize your dataset:  $Z = (x - \mu) / \text{std}$

Initial theta parameters is equal to zero. Learning rate is 0.01. Now, let's complete the following table:

#Iterations	Cost Function (Round please up to integer value)	Optimal Theta parameter Indicate here maximum theta value(Round please up to integer value)
n=100	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
n=1000	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
n=10000	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>

1234567890
Question **3**

Not yet answered

Marked out of 10.00

Suppose that you have the following dataset, with 3 input features, and 1 output feature. You're going to apply Logistic Regression algorithm with regularization.

Firstly it is necessary to apply normaization with the following formula:  $Z = (X - \mu) / \text{std}$ .

Initial theta parameters = 0.

#Iterations, lambda, learning rate	Cost function (rounded up to 2 digis after floating point)	Optimal theta parameter Indicate here maximum theta value (rounded up to 2 digis after floating point)
N=100, alpha = 1, lambda = 1	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
N=1000, alpha = 1, lambda = 10	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
N=10000, alpha = 2, lambda = 5	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>

After 10.000 iterations, alpha = 2, lambda = 5 and by setting threshold = 0.5, what is the number of zeros in the first 10 rows of prediction:

1234567890,

## Question 4

Not yet answered

Marked out of 10.00

Suppose that you're going to run neural network algorithm (see attached document):

$a5 = [ \square, \square ]$

$a4.max() = \square\square\square\square$  (rounded up to integer)

$a3.max() = \square\square\square$  (rounded up to integer)

$a2.min() = \square\square$  (rounded up to integer)

$a1.min() = \square$

General Conclusion:

Predicts image of cat

Predicts image of dog

## Question 5

Not yet answered

Marked out of 9.00

Suppose that you have the following dataset with 4 input features, 1 output variable (0,1 or 2). Your main task is to apply LogisticRegression algorithm and define precision, recall, accuracy and F-1 score for each class.

First 60% of dataset should be training set and last 40% test set.

Please round up to 3 digits after floating point. Drag-and-drop answers digit-by-digit.

List of necessary libraries:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
# Train a Logistic Regression classifier
clf_lr = LogisticRegression(random_state=42, max_iter=1000)
```

Accuracy:

F-1 score (class = 0):

F-1 score (class = 1):

F-1 score (class = 2):

◀ [FinalExam\\_7th\\_variant](#)

Jump to...