DSC 423: Data Analysis and Regression Assignment 01

Saransh Thakur

Id: 2020070

**Problem 1**

a) Take a simple random sample of 100 graduate students at DePaul University

b) Take a simple random sample of 100 graduate students studying Data Science.

The 100 graduate students studying Data science will have less standard deviation because we will have a specific group from where we need to choose students as well as textbooks; so, they will differ less. As they will differ less, they will have a less standard deviation.

Whereas in the case of 100 graduate students at DePaul University, we can choose students from any department, and they will have different textbooks as well. So, it will vary more. Hence, it will have more spread.

**Problem 2**

1. Mean= 28 and standard deviation =4

Text, letter

Description automatically generated

The percentage of students that are between 24 and 32 is 65%.

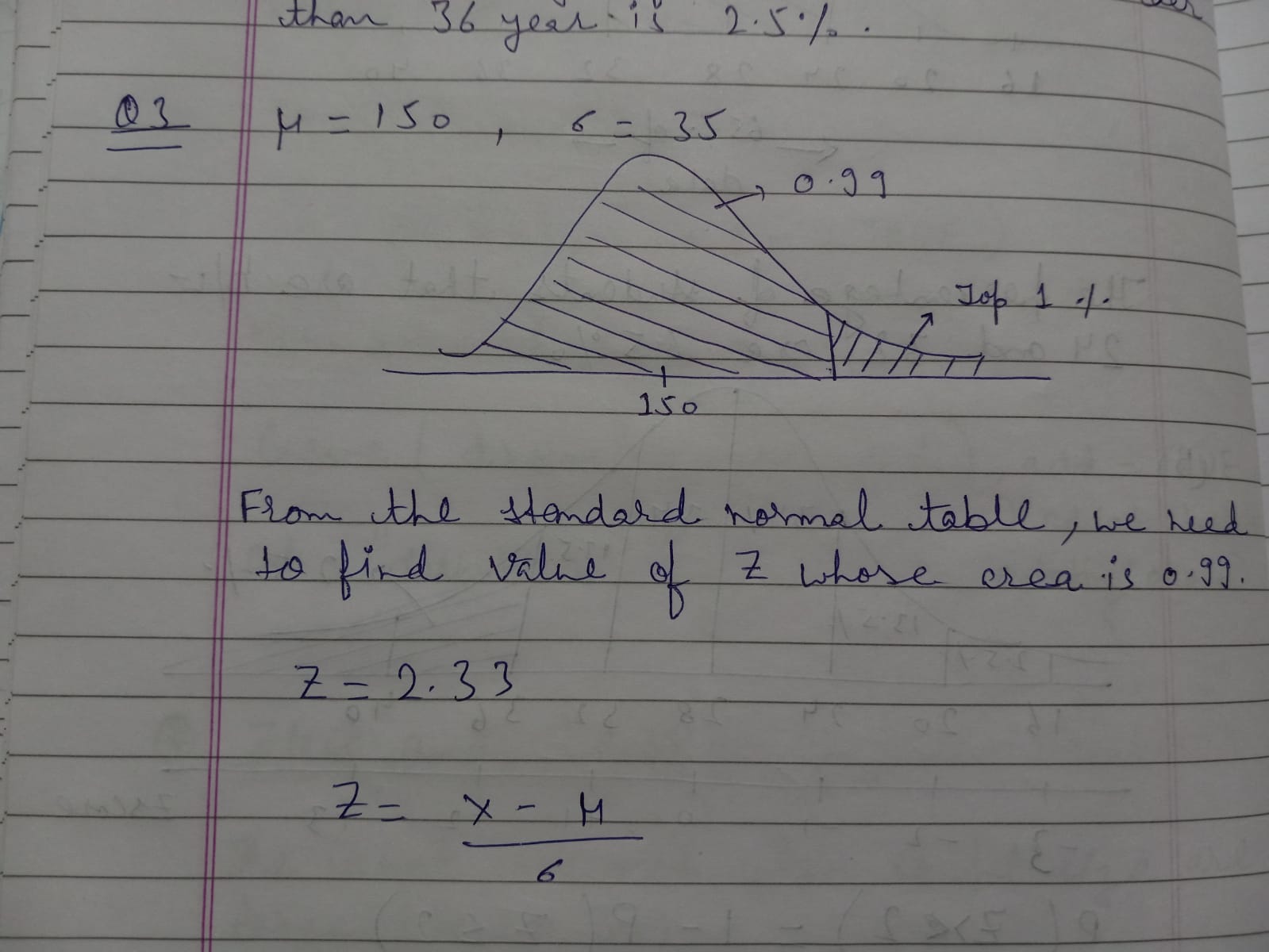
Text, letter

Description automatically generated

The percentage of students that are older than 36 years is 2.5%

**Problem 3**

Mean = 150, standard deviation= 35



From, the standard normal table, we need to find the value of Z whose area is 0.99 which is 2.33

Z is 2.33

Z = (x – mean)/ st.dev

2.33 = (x - 150)/35

**X = 231.55**

So, the top 1% monthly sale is 231.55 thousand dollars.

**Problem 4**

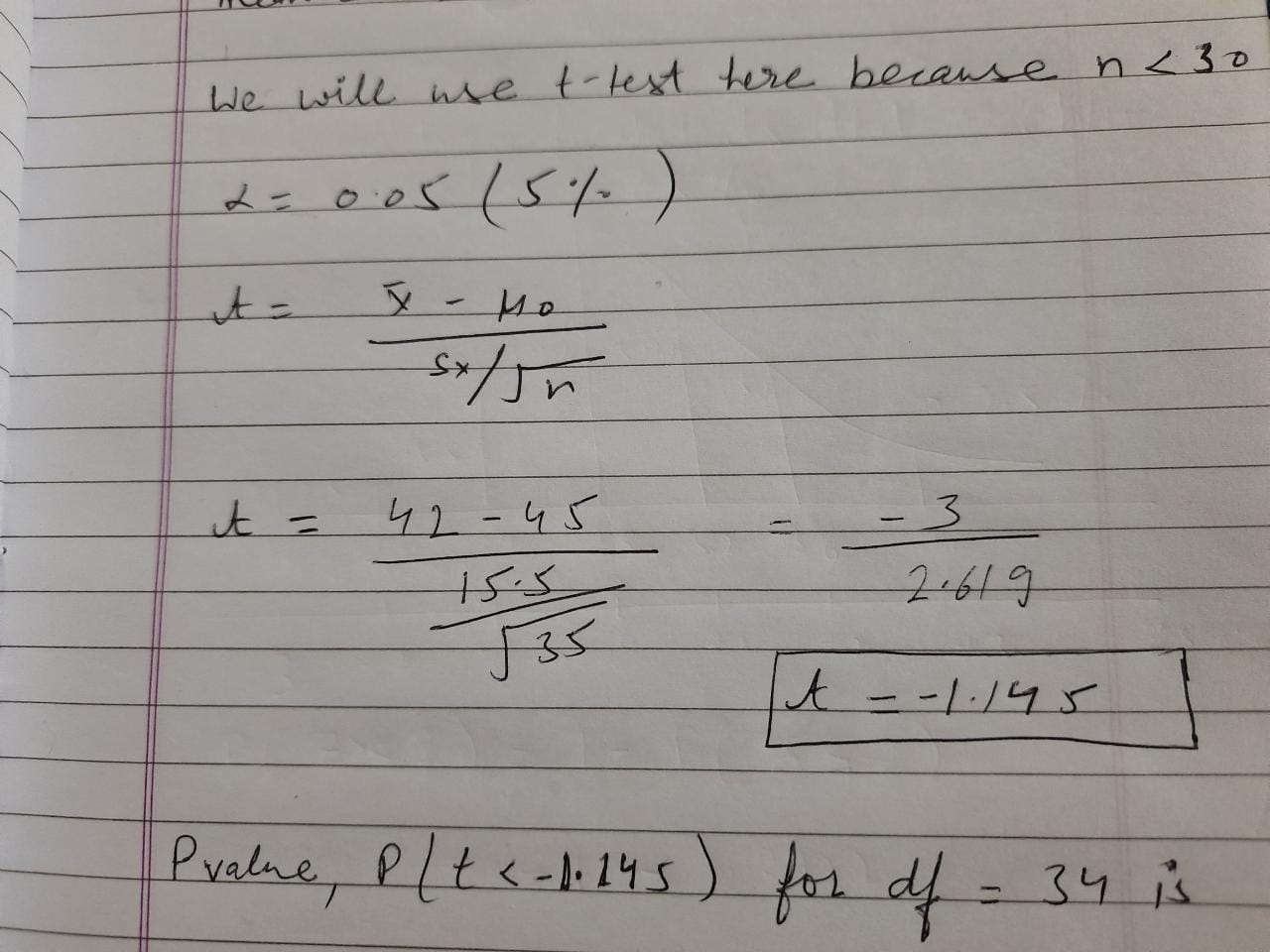
Ho (null hypothesis) = 45

H1 (Alternative hypothesis) < 45

n = 35, so df = n-1 = 34

mean = 42 and std.dev = 15.5

We will use t-test here because n < 30 and alpha (α) = 0.05 (5%)



P (t< 1.145) for df = 34 is between 0.15 and 0.10.

They both are greater than 5%. Hence, we fail to reject Ho (null hypothesis) which means there were on average 45 blocked intrusions per day.

"I have completed this work independently. The solutions given are entirely my work."

