



ACADGILD

SESSION 9: Statistical Inference

Assignment 1

Table of Contents

1. Introduction	3
2. Objective	3
3. Prerequisites	3
4. Associated Data Files	3
5. Problem Statement.....	3
7. Approximate Time to Complete Task	9

6.Expected Output	3
-------------------------	---

1. Introduction

This assignment will help you understand the concepts learnt in the session.

2. Objective

This assignment will test your skills on **theorems and tests** in R.

3. Prerequisites

Not applicable.

4. Associated Data Files

Not applicable.

5. Problem Statement

1. If Z is norm (mean = 0, sd = 1)

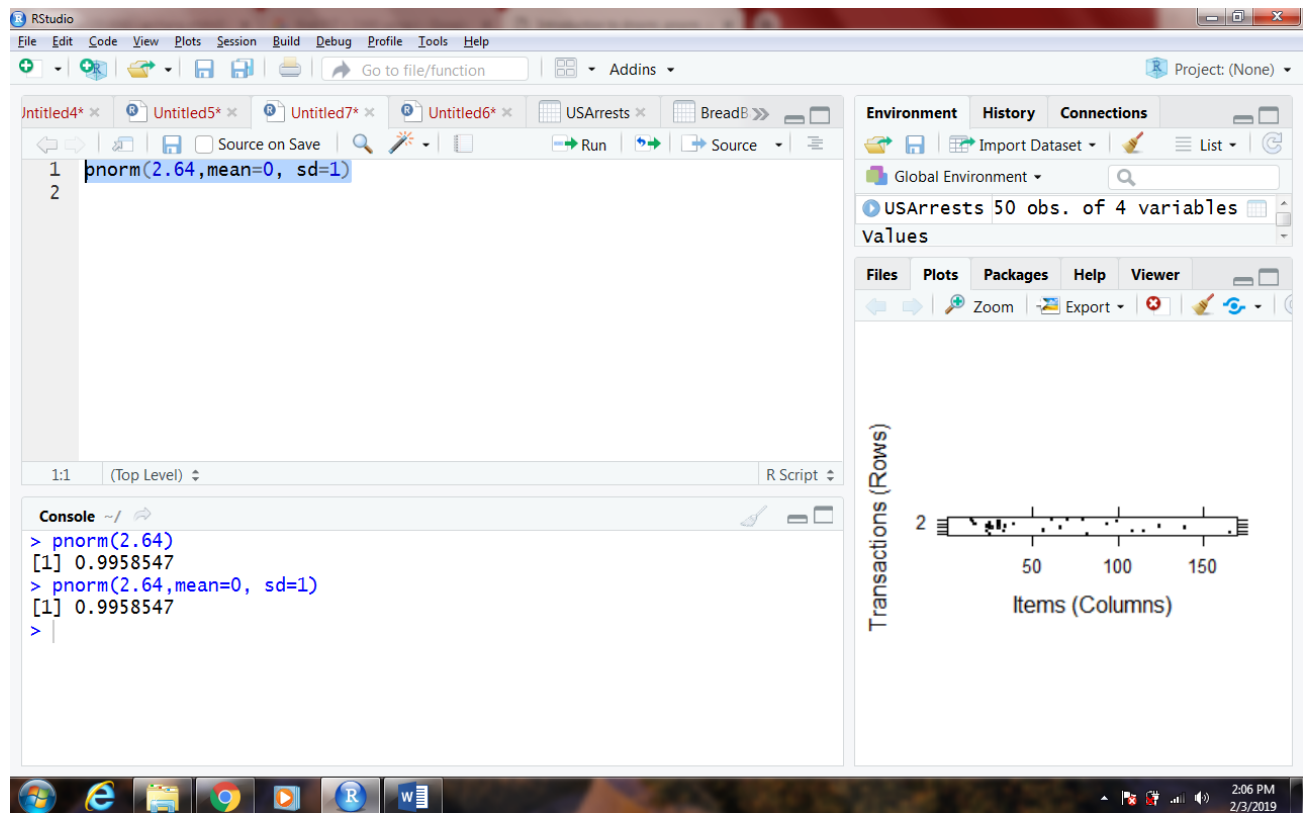
find $P(Z > 2.64)$

Ans =pnorm(2.64)

```
> pnorm(2.64)
[1] 0.9958547
```

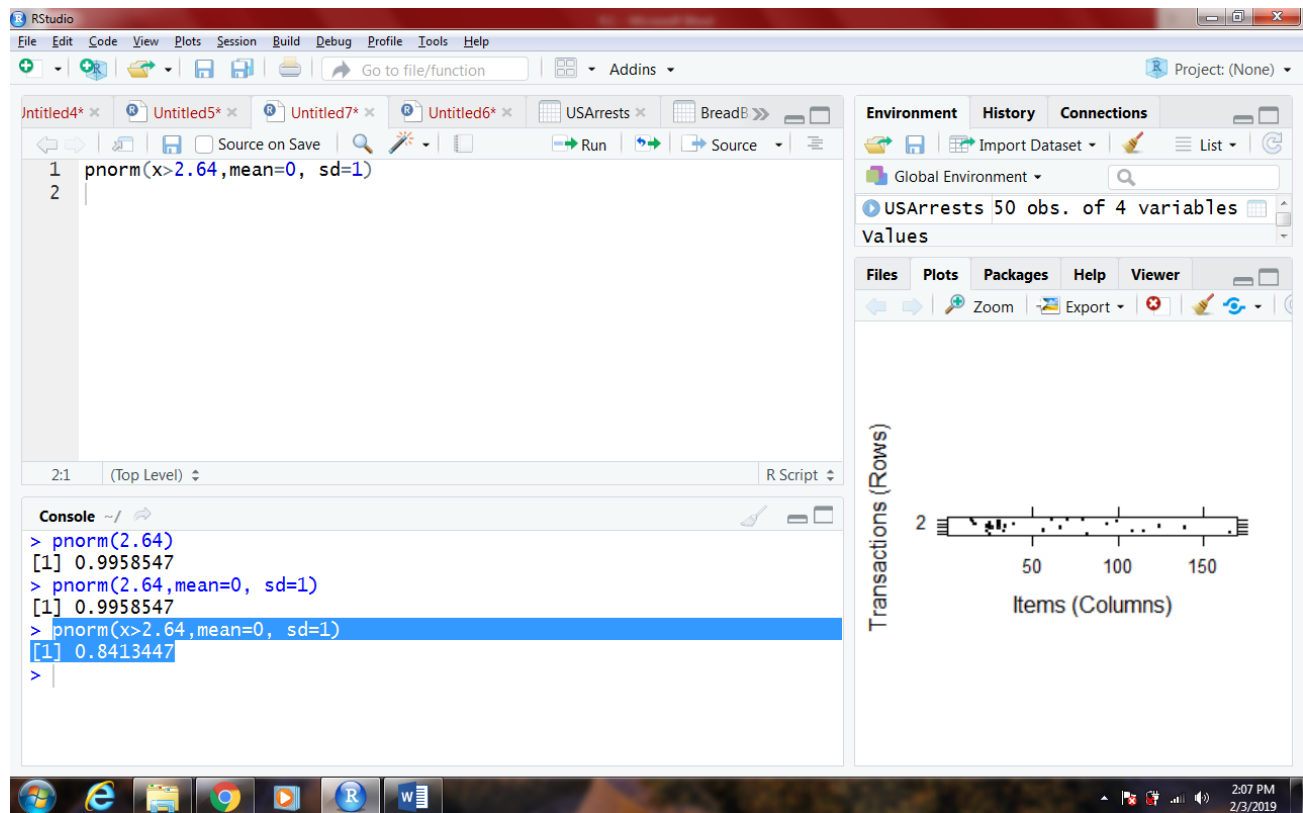
```
pnorm(2.64,mean=0, sd=1)
```

Data Analytics



```
pnorm(x>2.64, mean=0, sd=1)
[1] 0.8413447
```

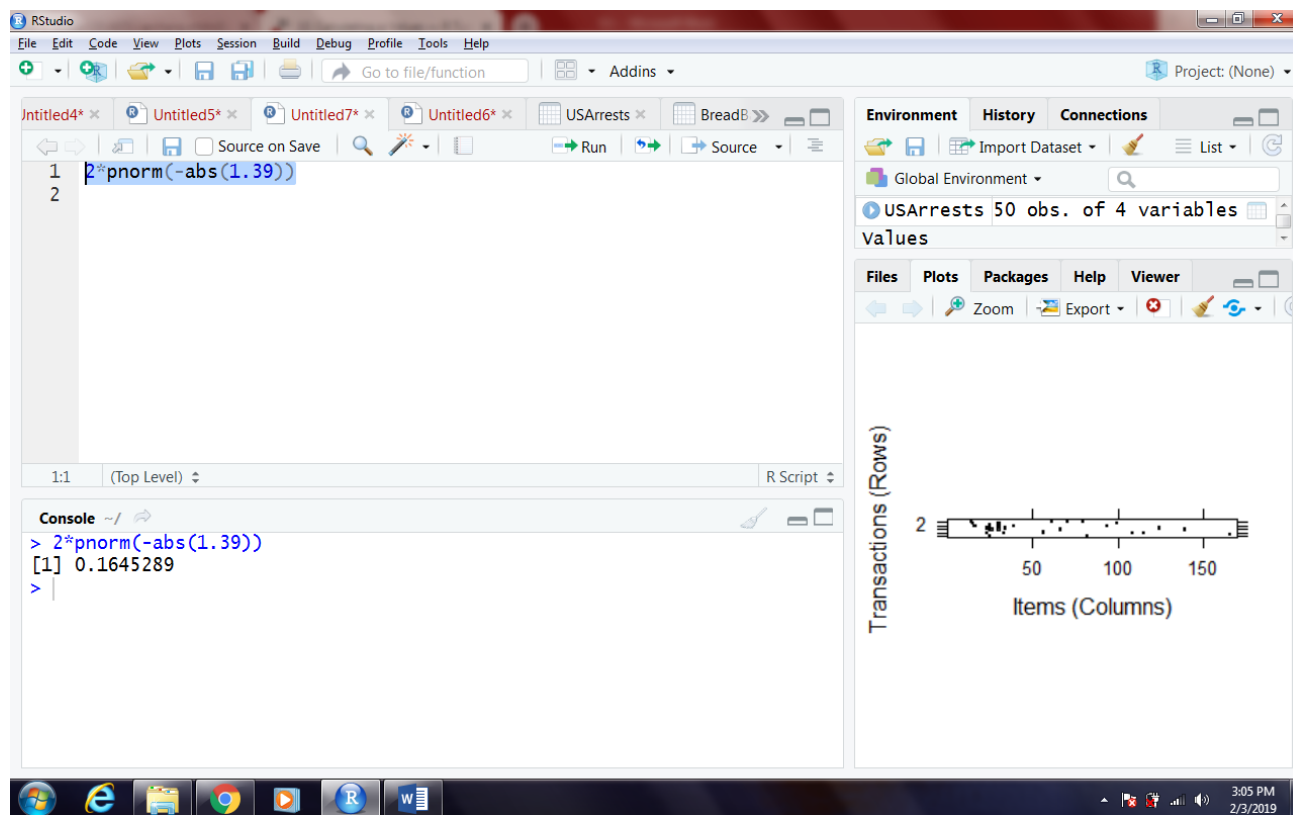
Data Analytics



find $P(|Z| > 1.39)$

```
2*pnorm(-abs(1.39))
```

Data Analytics



2. Suppose p = the proportion of students who are admitted to the graduate school of the University of California at Berkeley, and suppose that a public relation officer boasts that UCB has historically had a 40% acceptance rate for its graduate school. Consider the data stored in the table `UCBAdmissions` from 1973. Assuming these observations constituted a simple random sample, are

they consistent with the officer's claim, or do they provide evidence that the acceptance rate was significantly less than 40%?
Use an $\hat{\alpha} = 0.01$ significance level.

Ans- Our null hypothesis in this problem is $H_0 : p = 0.4$ and the alternative hypothesis is $H_1 : p < 0.4$. We reject the null hypothesis if \hat{p} is too small, that is, if

$$\hat{p} - 0.4 \sqrt{0.4(1 - 0.4)/n} < -z_{\alpha},$$

where $\alpha = 0.01$ and $-z_{0.01}$ is

`qnorm(0.99)`

```
[1] 2.326348
```

Our only remaining task is to find the value of the test statistic and see where it falls relative to the critical value. We can find the number of people admitted and not admitted to the UCB graduate school with the following.

```
A<-as.data.frame(UCBAdmissions)
head(A)
```

n	Admit	Gender	Dept	Freq
1	1	Admitted	Male	A 512
2	2	Rejected	Male	A 313
3	3	Admitted	Female	A 89
4	4	Rejected	Female	A 19
5	5	Admitted	Male	B 353
6	6	Rejected	Male	B 207

>

```
xtabs(Freq~Admit,data=UCBAdmissions)
```

Admit
Admitted Rejected
1755 2771

>

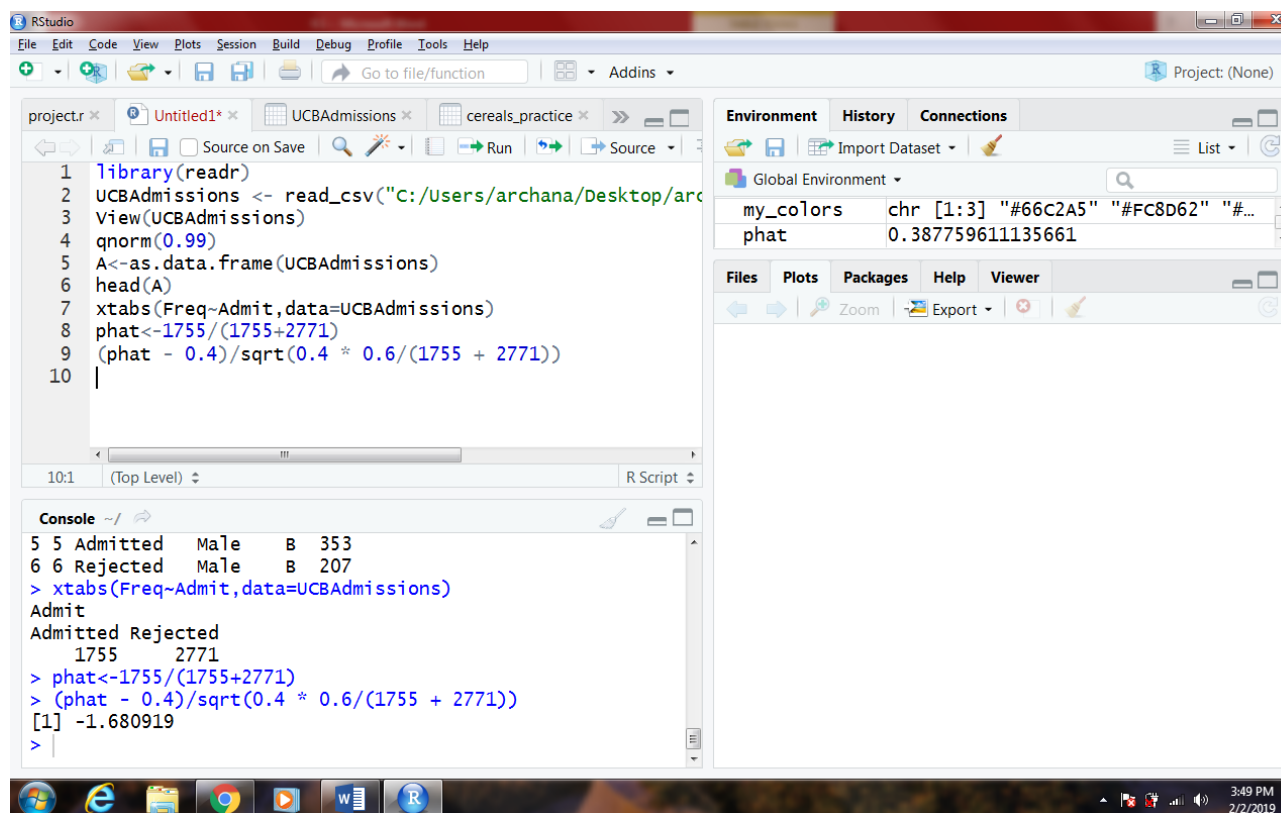
Now we calculate the value of the test statistic.

```
phat<-1755/(1755+2771)
(phat - 0.4)/sqrt(0.4 * 0.6/(1755 + 2771))
```

Data Analytics

```
[1] -1.680919
```

Our test statistic is not less than -2.32 , so it does not fall into the critical region. Therefore, we fail to reject the null hypothesis that the true proportion of students admitted to graduate school is less than 40% and say that the observed data are consistent with the officer's claim at the $\alpha = 0.01$ significance level.



6. Expected Format

1. R file should be submitted where applicable.
2. R file should be in PDF or in .r format
3. Proper screenshots of the outputs should be submitted as well
4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

7. Approximate Time to Complete Task

40 mins.