Task1: To prove if U.S managers rate significantly less than 4.3

Ho(Null)-The U.S managers are rated 4.3 on average

H1(Alternate)-The U.S managers are rated less than 4.3

Tail selection = Left tailed selection test

In [1]:
```
import numpy as np
import statistics as st
import matplotlib.pyplot as plt
from scipy.stats import norm
import pandas as pd
```

In [2]:
```
mean_p=4.3
S=0.574
alpha=0.1
samp_data=[int(i) for i in ("3 4 5 5 4 5 5 4 4 4 4 4 4 4 4 5 4 4 4 3 3 3 4 3 5 4 4 5 4 4 4 5".split())]
mean_s=st.mean(samp_data)
n_s=len(samp_data)

#Calculating z_statistics, p_value and Z_critical

z_critical=norm.ppf(1-alpha/2)

z_stats=(mean_s-mean_p)/(S/(n_s)**0.5)

p_value=norm.sf(abs(z_stats))*2

print("The z_statistics = ",z_stats)
print("The z_critical = ",z_critical)
print("The p_value = ",p_value)
print("The alpha = ",alpha)
```

```
The z_statistics =  -2.032624022574569
The z_critical =  1.6448536269514722
The p_value =  0.042090522594647045
The alpha =  0.1
```

Rules of rejection of null hypothesis:

1) The p_value < alpha - Which is True rejection of Ho

2) The z_statistics must not be lesser than or greater than z_critical in our case it is lesser i.e z_statistics< -z_critical

Thus we reject the null hypothesis and accept the alternate hypothesis that U.S managers are rated less than 4.3 or greater than 4.3

Business decision : The ratings of managers can be improved by trainings and cultural events and recruiting more employers might reduce stress on them

Task2: Health Drink price survey

In [4]:
```
city_p=pd.read_csv(r"K:\Desktop\NIIT\tables\DS1_C5_S6_CityPrice_Data_Challenge.csv")
city_p.Florida.fillna(np.mean(city_p.Florida),inplace=True)

mean_c=city_p["California"].mean()
mean_f=city_p["Florida"].mean()

n1=33
n2=30

var_c=0.028
var_f=0.015
sc=var_c**0.5
sf=var_f**0.5
```

Ho(Null)= The mean of entire price is same

H1(Allternate)= The mean of entire price is increased or decreased and not the same

here mu!=0 (Two tail test)

In [5]:
```
# b)
alpha=0.01

z_critical=norm.ppf(1-alpha/2)

print("The z_critical value = ",z_critical)

z_statistics = (mean_f-mean_c)/(sc/(n2)+sf/(n1))**0.5
print("The z_statistics value is = ",z_statistics)

p_val=norm.sf(abs(z_statistics))*2

print("The p value = ",p_val)

# calculating confidence interval mean difference
ls=(mean_c - mean_f) - 2.58*((var_c/n1 + var_f/n2))

rs=(mean_c - mean_f) + 2.58*((var_c/n1 + var_f/n2))
print("The difference in price with 99% confidence level = ",rs-ls)

```

```
The z_critical value =  2.5758293035489004
The z_statistics value is =  -1.0444780056077538
The p value =  0.2962642899686514
The difference in price with 99% confidence level =  0.006958181818181808
```

Rules of rejection of null hypothesis :

1) The p_value < alpha - Which is False rejection of Ho

2) The z_statistics must not be lesser than or greater than z_critical in our case it is lesser i.e z_statistics< -z_critical

In our case both the rules apply and thus null hypothesis is True and we can reject alternate hypothesis

Business decision:

The price of health drink has not changed since we have rejected the alternate hypothesis

Ho(Null)- The average age of female trader is 44

H1(Alternate) - The age is higher

In [6]:
```
pop_mean=44
ch_mean=45.1
n1=68
S=8.7

alpha=0.05

z_critical=norm.ppf(1-alpha/2)

print("The z_critical value = ",z_critical)

z_statistics = (ch_mean-pop_mean)/(S/(n1)**0.5)
print("The z_statistics value is = ",z_statistics)

p_val=norm.sf(abs(z_statistics))*2

print("The p value = ",p_val)
```

```
The z_critical value =  1.959963984540054
The z_statistics value is =  1.042624411075732
The p value =  0.29712227883324194
```

In [7]:
```
if(p_val<alpha and z_critical<z_statistics):
    print("Reject Null Hypothesis")
else:
    print("Accept Null hypothesis")
```

```
Accept Null hypothesis
```

Decision on situation:

Since we have accepted the null hypothesis that is the average age of female trader in stock exchange is unchanged thus 44.