**Assignment z-test**

Date: 05/05/2020 Name: D.Saravanan

1. A sample of 400 male students is found to have a mean height of 171.38 cm. Can it be reasonably regarded as a sample from a large population with mean height 171.17 cm and standard deviation 3.30 cm.

   **Solution:**

   Given that $n = 400$, $\bar{x} = 171.38$ cm, $\mu = 171.17$ cm, $\sigma = 3.30$ cm

   **Null hypothesis:** $H_0 : \mu = 171.17$, No difference between sample mean and hypothetical population mean
   **Alternative hypothesis:** $H_1 : \mu \neq 171.17$ (two tailed test)

   $z$-statistic:

   $$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{171.38 - 171.17}{3.30/\sqrt{400}} = 1.27273$$

   At $5\%$ significance level the tabulated value for $z_{\alpha/2}$ is $1.96$

   **Conclusion:** $|z| < z_{\alpha/2}$, we accept the Null hypothesis. That is there is no significant difference between the sample mean and the population mean.

   Standard Error:

   $$S.E = \frac{\sigma}{\sqrt{n}} = \frac{3.30}{\sqrt{400}} = 0.165$$

   Confidence Interval: $95\%$ fiducial limits
   The confidence interval for the population mean is given by

   $$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
   $$171.38 \pm 1.96\frac{3.30}{\sqrt{400}} = 171.38 \pm 0.3234$$
   $$171.0566 < \mu < 171.7034$$

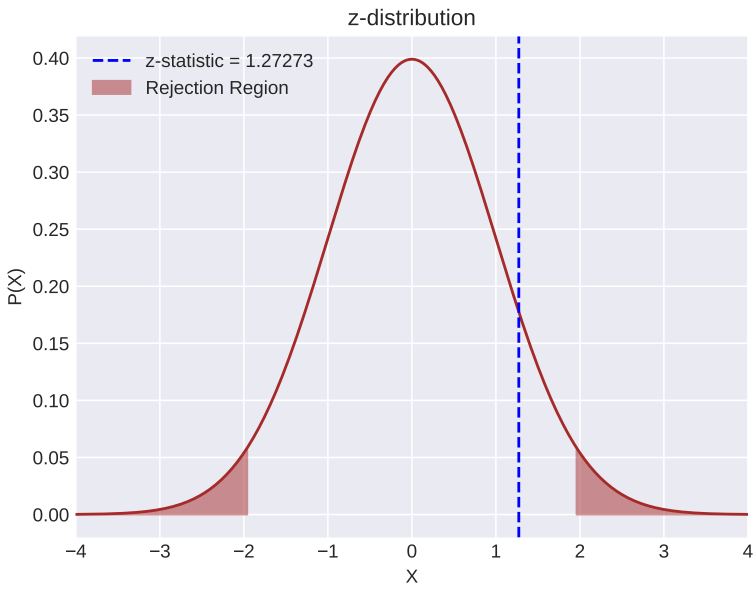
   Program:

```
#!/usr/bin/env python3
import numpy as np
from scipy.stats import norm
import matplotlib.pyplot as plt
plt.style.use('seaborn-darkgrid')

print("__________z-test__________")

n = float(input("\nSample length: "))
s_mean = float(input("Sample mean: "))
p_mean = float(input("Population mean: "))
p_stdv = float(input("Population stand_dev: "))
alpha = float(input("Level of significance: "))

# calculation of z-statistic
zstatistic = (s_mean - p_mean)/(p_stdv/np.sqrt(n))
```

```
print("\nz-statistic: {:.5f}".format(zstatistic))

# calculation of critical values
zcritical_l = norm.ppf(q = alpha/2)
zcritical_u = -zcritical_l
print("\nCritical values are {:.5f}, {:.5f}".format(zcritical_l, zcritical_u))

# decision making: z-statistic and critical values
if zstatistic < zcritical_l or zstatistic > zcritical_u:
    print("Reject the Null hypothesis.")
else: print("Fail to reject the Null hypothesis.")

# calculation of p-value:
if zstatistic < 0: pvalue = 2*norm.cdf(zstatistic)
else: pvalue = 2*(1 - norm.cdf(zstatistic))
print("\np-value: {:.5f}".format(pvalue))

# decision making: p-value and level of significance
if pvalue < alpha: print("Reject the Null hypothesis.")
else: print("Fail to reject the Null hypothesis.")

# standard error
std_err = p_stdv/np.sqrt(n)
print("\nStandard Error: {:.5f}".format(std_err))

# confidence interval
cnf_int = s_mean + std_err * np.array([zcritical_l, zcritical_u])
print("Confidence Interval: {}".format(cnf_int))

x1=np.linspace(-10,zcritical_l,1000); y1=norm.pdf(x1)
x2=np.linspace(zcritical_u,10,1000);  y2=norm.pdf(x2)
x3=np.linspace(-4,4,1000); y3=norm.pdf(x3)
ax=plt.figure().add_subplot(111); ax.plot(x3,y3,color='brown')
ax.fill_between(x1,y1,0,alpha=0.5,color='brown',label='Rejection Region')
ax.fill_between(x2,y2,0,alpha=0.5,color='brown'); ax.set(xlabel='X',ylabel='P(X)')
ax.axvline(x=zstatistic,ls='--',c='b',label='z-statistic = {:.5f}'.format(zstatistic))
ax.set(xlim=[-4,4],title="z-distribution"); plt.legend();
plt.savefig('zplot1.pdf',dpi=72,bbox_inches='tight'); plt.show()
```

Output:

```
__________z-test__________

Sample length: 400
Sample mean: 171.38
Population mean: 171.17
Population stand_dev: 3.30
Level of significance: 0.05

z-statistic: 1.27273

Critical values are -1.95996, 1.95996
Fail to reject the Null hypothesis.

p-value: 0.20311
Fail to reject the Null hypothesis.

Standard Error: 0.16500
Confidence Interval: [171.05660594 171.70339406]
```




2. A sample of 900 items has mean 3.4 and standard deviation 2.61 Can the sample be regarded as drawn from a population with mean 3.25 at 1 percent level of significance.

**Solution:**

Given that $n = 900$, $\bar{x} = 3.4$, $s = 2.61$, $\mu = 3.25$

**Null hypothesis:** $H_0 : \mu = 3.25$, No difference between sample mean and hypothetical population mean
**Alternative hypothesis:** $H_1 : \mu \neq 3.25$ (two tailed test)

$z$-statistic:

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{3.4 - 3.25}{2.61/\sqrt{900}} = 1.72414$$

At $1\%$ significance level the tabulated value for $z_{\alpha/2}$ is 2.576

**Conclusion:** $|z| < z_{\alpha/2}$, we accept the Null hypothesis. The sample can be regarded as drawn from a population with mean 3.25 at 1 percent level of significance.

Standard Error:

$$S.E = \frac{s}{\sqrt{n}} = \frac{2.61}{\sqrt{900}} = 0.087$$

Confidence Interval: $99\%$ fiducial limits
The confidence interval for the population mean is given by

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$
$$3.4 \pm 2.576 \frac{2.61}{\sqrt{900}} = 3.4 \pm 0.2241$$
$$3.1759 < \mu < 3.6241$$

Program:

```
#!/usr/bin/env python3
import numpy as np
from scipy.stats import norm
import matplotlib.pyplot as plt
plt.style.use('seaborn-darkgrid')

print("__________z-test__________")

# sample length
n = 900

# sample mean
s_mean = 3.4

# sample standard deviation
s_stdv = 2.61

# population mean
p_mean = 3.25

# level of significance
alpha = 0.01

# calculation of z-statistic
zstatistic = (s_mean - p_mean)/(s_stdv/np.sqrt(n))
print("\nz-statistic: {:.5f}".format(zstatistic))

# calculation of critical values
zcritical_l = norm.ppf(q = alpha/2)
zcritical_u = -zcritical_l
print("\nCritical values are {:.5f}, {:.5f}".format(zcritical_l, zcritical_u))

# decision making: z-statistic and critical values
if zstatistic < zcritical_l or zstatistic > zcritical_u:
    print("Reject the Null hypothesis.")
else: print("Fail to reject the Null hypothesis.")

# calculation of p-value:
if zstatistic < 0: pvalue = 2*norm.cdf(zstatistic)
else: pvalue = 2*(1 - norm.cdf(zstatistic))
print("\np-value: {:.5f}".format(pvalue))

# decision making: p-value and level of significance
if pvalue < alpha: print("Reject the Null hypothesis.")
else: print("Fail to reject the Null hypothesis.")

# standard error
std_err = s_stdv/np.sqrt(n)
print("\nStandard Error: {:.5f}".format(std_err))

# confidence interval
cnf_int = s_mean + std_err * np.array([zcritical_l, zcritical_u])
print("Confidence Interval: {}".format(cnf_int))

x1=np.linspace(-10,zcritical_l,1000); y1=norm.pdf(x1)
x2=np.linspace(zcritical_u,10,1000);  y2=norm.pdf(x2)
x3=np.linspace(-4,4,1000); y3=norm.pdf(x3)
ax=plt.figure().add_subplot(111); ax.plot(x3,y3,color='brown')
ax.fill_between(x1,y1,0,alpha=0.5,color='brown',label='Rejection Region')
ax.fill_between(x2,y2,0,alpha=0.5,color='brown'); ax.set(xlabel='X',ylabel='P(X)')
ax.axvline(x=zstatistic,ls='--',c='b',label='z-statistic = {:.5f}'.format(zstatistic))
ax.set(xlim=[-4,4],title="z-distribution"); plt.legend();
plt.savefig('zplot2.pdf',dpi=72,bbox_inches='tight'); plt.show()
```

Output:

```
__________z-test__________

z-statistic: 1.72414

Critical values are -2.57583, 2.57583
Fail to reject the Null hypothesis.

p-value: 0.08468
Fail to reject the Null hypothesis.

Standard Error: 0.08700
Confidence Interval: [3.17590285 3.62409715]
```