In [2]:

```
# 1

# 1. Yes, because values in both are statement are about population and have equal values and inequality in alternate hypothesis.
# 2. No, because null hypothesis has greater than claim and alternate hypothesis has equality claim.
# 3. Yes, because values in both are statement are about population, null is equal and alternate is claiming inequality
# 4. No, because values in both hypothesis is different and has equal sign.
# 5. No, because hypothesis are always statements about population or distribution and not about sample
```

In [41]:

```
# 2

from scipy.stats import norm
import numpy as np
import math
import scipy.stats as stats
from scipy.stats import chi2_contingency

p_mean = 52
p_std = 4.50
n = 100
sample_mean = 52.80

SE = p_std/n**0.5
Z = (sample_mean-p_mean)/SE
print(f"z_score is:{Z}")
alpha=0.05  #test_significance
print(f"Critical region: {norm.ppf(alpha/2)}, {-norm.ppf(alpha/2)}")
```

```
z_score is:1.7777777777777715
Critical region: -1.9599639845400545, 1.9599639845400545
```

In [40]:

```
# 3

p_mean = 34
p_std = 8
n = 50
sample_mean = 32.5

SE = p_std/n**0.5   #standard Error
Z = (sample_mean-p_mean)/SE
print(f"z_score is:{Z}")
alpha=0.01
print(f"Critical region: {norm.ppf(alpha/2)}, {-norm.ppf(alpha/2)}")
```

```
z_score is:-1.3258252147247767
Critical region: -2.575829303548901, 2.575829303548901
```

In [42]:

```
# 4
data = [1008, 812, 1117, 1323, 1308, 1415, 831, 1021, 1287, 851, 930, 730, 699, 872, 913, 944, 954, 987, 1695, 995, 1003, 994]
p_mean =1135
sample_std = np.std(data)
n=22
sample_mean = np.sum(data,axis=0)/len(data)
SE = sample_std/n**0.5
alpha = 0.5
test_1 = (sample_mean-p_mean)/SE
print(f"t_score is{test_1}")
print(f"Critical Region: {stats.t.ppf((alpha/2),df=21)} {stats.t.ppf(1-(alpha/2),df=21)}")
```

```
t_score is-2.070747228595759
Critical Region: -0.6863519891164291 0.6863519891164291
```

In [13]:

```
# 5

p_mean = 48432
p_std = 2000
n =400
sample_mean =48574

SE = p_std/n**0.5
Z = (sample_mean-p_mean)/SE
alpha=0.05
print(f"Critical region: {norm.ppf(alpha/2)} {-norm.ppf(alpha/2)}")
```

```
Critical region: -1.9599639845400545 1.9599639845400545
```

In [12]:

```
# 6

p_mean =32.28
n=19
sample_mean =31.67
sample_std =1.29
alpha =0.05

SE=sample_std/(n**0.5)
t=(sample_mean-p_mean)/SE
print(f"t_score is {round((t),1)}")
print(f"Critical region: {round(stats.t.ppf((alpha/2),df=18),1)} {-round(stats.t.ppf((alpha/2),df=18),1)}")
```

```
t_score is -2.1
Critical region: -2.1 2.1
```

In [20]:

```
# 8

n = 16
p_mean = 10
sample_mean =12
sample_std =1.5

SE = sample_std/(n**0.5)
t = (sample_mean-p_mean)/SE
print(f"t_score: {round((t),1)}")
```

```
t_score: 5.3
```

In [21]:

```
# 9

n= 16
alpha=(1-0.99)/2
print(f"t_score: {stats.t.ppf(1-alpha,df=15)}")
```

```
t_score: 2.946712883338615
```

In [23]:

```
# 10

n=25
std=4
mean=60
alpha=(1-0.95)/2
t_score=stats.t.ppf(1-alpha,df=24)
print(f"Range: {mean+t_score*(std/(n**0.5))} {mean-t_score*(std/(n**0.5))}")
p = stats.t.cdf(0.1,df=24)-stats.t.cdf(-0.05,df=24)
print(f"probability that (−t0.05 <t<t0.10) is {p}")
```

```
Range: 61.651118849302414 58.348881150697586
probability that (−t0.05 <t<t0.10) is 0.05914441613731247
```

In [43]:

```
# 11

n1 = 1200
x1 = 452
s1 = 212
n2 = 800
x2 = 523
s2 = 185
s_1=s1**2
s_2=s2**2
alpha=0.05
se=((s_1/n1)+(s_2/n2))**0.5
z_score=(x1-x2)/se
print(f"z_score: {z_score}")
print(f"Critical region: {norm.ppf(alpha/2)} {-norm.ppf(alpha/2)}")
# Conclusion- reject null hypothesis, number of people travelling from Bangalore to Chennai is different from the number
# of people travelling from Bangalore to Hosur in a week
```

```
z_score: -7.926428526759299
Critical region: -1.9599639845400545 1.9599639845400545
```

In [44]:

```
# 12

n1 = 100
x1 = 308
s1 = 84
n2 = 100
x2 = 254
s2 = 67
s_1=s1**2
s_2=s2**2
alpha=0.05
SE=((s_1/n1)+(s_2/n2))**0.5
z_score=(x1-x2)/SE
print(f"z_score: {z_score}")
print(f"Critical region: {norm.ppf(alpha/2)} {-norm.ppf(alpha/2)}")
# Conclusion- reject null hypothesis, number of people preferring Duracell battery is different from the number of
# people preferring Energizer battery
```

```
z_score: 5.025702668336442
Critical region: -1.9599639845400545 1.9599639845400545
```

In [32]:

```
# 13

n1 = 14
x1 = 0.317
s1 = 0.12
n2 = 9
x2 = 0.21
s2 = 0.11
s_1=s1**2
s_2=s2**2
s=((n1-1)*s_1)+((n2-1)*s_2)
n=(n1+n2-2)
se=(s/n)**0.5
n_1=((1/n1)+(1/n2))**0.5
t_score=(x1-x2)/se*n_1

print(f"t_score: {t_score}")
print(f"Critical Region: {stats.t.ppf(1-0.05,df=n)}")
# Conclusion- accept null hypothesis, average price do not increase
```

```
t_score: 0.3931089218182991
Critical Region: 1.7207429028118775
```

In [34]:

```
# 14

n1 = 15
x1 = 6598
s1 = 844
n2 = 12
x2 = 6870
s2 = 669
s_1=s1**2
s_2=s2**2

s=((n1-1)*s_1)+((n2-1)*s_2)
n=(n1+n2-2)
se=(s/n)**0.5
n_1=((1/n1)+(1/n2))**0.5
t_score=(x1-x2)/se*n_1

print(f"t_score: {t_score}")
print(f"Critical Region: {stats.t.ppf(0.05,df=n)}")
# Conclusion- accept null hypothesis, average price remains same
```

```
t_score: -0.1364745051598569
Critical Region: -1.708140761251899
```

In [45]:

```
# 15

n1 = 1000
x1 = 53
p1 = 0.53
n2 = 100
x2 = 43
p2= 0.53
p=(x1+x2)/(n1+n2)

n=(1/n1)+(1/n2)
p_1=p*(1-p)
z=(p1-p2)/((p_1*n)**0.5)
print(f"z_score: {z}")
print(f"Critical region: {norm.ppf(0.05)}")
# Conclusion- accept null hypothesis
```

```
z_score: 0.0
Critical region: -1.6448536269514729
```

In [46]:

```
# 16

n1 = 300
x1 = 120
p1 = 0.40
n2 = 700
x2 = 140
p2= 0.20
p=(x1+x2)/(n1+n2)

n=(1/n1)+(1/n2)
p_1=p*(1-p)
Z=(p1-p2-0.1)/((p_1*n)**0.5)
print(f"z_score: {Z}")
print(f"Critical region: {-norm.ppf(0.05)}")
# Conclusion- accept null hypothesis
```

```
z_score: 3.303749523611152
Critical region: 1.6448536269514729
```

In [53]:

```
# 17

f_obs= [16, 20, 25, 14, 29, 28]
f_exp= [22,22,22,22,22,22]
result=stats.chisquare(f_obs,f_exp)
print(f"Chi square value: {result[0]} and p-value: {result[1]}")
# Conclusion- Dice is unbiased
```

```
Chi square value: 9.0 and p-value: 0.1090641579497725
```

In [56]:

```
# 19

obs=[41,19,24,16]
exp=[25,25,25,25]
result=stats.chisquare(obs,exp)
print(f"Chi Square value: {result[0]}")
print(f"Critical region with 3df and alpha=0.05: 7.82")
# Conclusion- Reject null hypothesis, all candidates are not equally popular
```

```
Chi Square value: 14.959999999999999
Critical region with 3df and alpha=0.05: 7.82
```

In [57]:

```
# 20

obs=([[18,22,20],[2,28,40],[20,10,40]])
result=chi2_contingency(obs)
print(f"Chi Square value: {result[0]}")
print(f"Critical region with 4df and alpha=0.001: 18.47")
# Conclusion- Reject null hypothesis, there is significant relationship between age and photograph preference
```

```
Chi Square value: 29.603174603174608
Critical region with 4df and alpha=0.001: 18.47
```

In [59]:

```
# 21

obs=np.array([[18,40],[32,10]])
result=chi2_contingency(obs)
print(f"Chi Square value: {result[0]}")
print(f"Critical region with 1df and alpha=0.001: 10.83")
# Conclusion- reject null hypothesis
```

```
Chi Square value: 18.10344827586207
Critical region with 1df and alpha=0.001: 10.83
```

In [61]:

```
# 22

obs=([[12,32],[22,14],[9,6]])
result=chi2_contingency(obs)
print(f"Chi Square value: {result[0]}")
print(f"Critical region with 2df and alpha=0.001: 13.82")
# Conclusion- accept null hypothesis, there is no relationship between height and leadership qualities
```

```
Chi Square value: 10.712198008709638
Critical region with 2df and alpha=0.001: 13.82
```

In [63]:

```
# 23

obs = np.array([[679,103,114], [63,10,20],[42,18,25]])
result=chi2_contingency(obs)
print(f"Chi Square value: {result[0]}")
print(f"Critical region with 4df and alpha=0.001: 18.47")
# Conclusion- reject null hypothesis, there is relationship between martial status and employment status")
```

```
Chi Square value: 31.61310319407798
Critical region with 4df and alpha=0.001: 18.47
```

In [ ]: