**Topic 7: Inference for the Proportion Solutions**

**Q1**

a) 

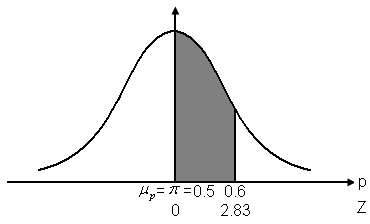
b) Standard error of the proportion,

**Q2**

1.  An individual has no ability to distinguish between the two brands.

  = proportion of students which can distinguish the brand = 0.5

n = 200 > 30, n=200(0.5)= 100 > 5, n(1-)=200(1-0.5)=100 > 5

So sampling distribution of p is approximately normal

 => 

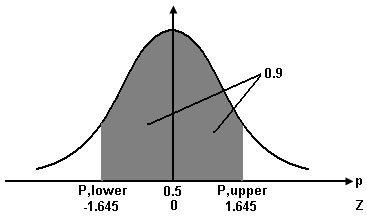










b) 



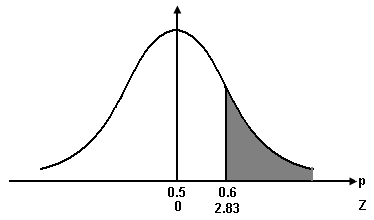
For symmetric distribution of probability (0.9) on both sides,

 and 

Since  and ,

|  |  |
| --- | --- |
|  |  |

Hence, 90% of the sample proportion will be between 0.4418 and 0.5582.

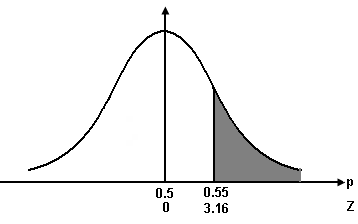
c) 

d) Case 1: 



= 1- 0.9977

= 0.0023



Case 2: 



= 1- 0.99921

= 0.00079

More than 60% correct identification in a sample of 200 is more likely to occur than more than 55% correct in a sample of 1000.

**Q3**

1. n=500>30, np=500(0.27) =135 >5, n(1-p)=500(1-0.27)=365>5

 Sampling distribution of *p* is approximately normal.

, , 

95% confidence interval for :

== [0.2311,0.3089]

We are 95% confident that the population proportion of small business owners who never check in with the office when on vacation is estimated to be between 0.2311 and 0.3089.

1. n=1000>30, np=1000(0.27) =270 >5, n(1-p)=1000(1-0.27)=730 >5

 sampling distribution of *p* is approximately normal.

, , 

95% confidence interval for :

== [0.2425,0.2975]

We are 95% confident that the population proportion of small business owners who never check in with the office when on vacation is estimated to be between 0.2425 and 0.2975.

1. The larger the sample size, the narrower is the confidence interval holding everything constant.

**Q4**

1. n=658>30, np=658(0.3799) =250 >5, n(1-p)= 658(1-0. 3799)=408 >5

 sampling distribution of *p* is approximately normal.

, , 

95% confidence interval for :

==[0.3429,0.4170]

We are 95% confident that the population proportion of CEOs whose greatest concern was sustained and steady top-line growth is estimated to be between 0.3429 and 0.4170.

1. Replace  by *p*

= 9049.92 9050 (round up)

**Q5**

n = ==860.5

If information is not available, let p=0.5

n’ = = 1067.11068

**Q6**

a) H: π0.95

H: π < 0.95



p-value=P(Z ≤ -1.69)=0.0455

Since p-value=0.0455 < 0.1=

Ho is rejected, and we conclude that there is not sufficient evidence to support that manager’s claim. i.e. Less than 95% of Pizza Delight’s orders are delivered within 15 minutes of the time when order is placed at the 0.1 level of significance.

b) The decision Rule is given as follows:

Ho is rejected if the given significance level;

Ho is not rejected if the given significance level 

**Q7**

1. Let  be the population proportion of customers who select products and then cancel their transaction

H0: 

H1: 

n =500> 30, np=210>5, n(1-p)=290>5

Sampling distribution of p is approximately normal

Reject H0 if Z < −2.33



Since Z =-3.577 < -2.33

Reject H0 at 

There is sufficient evidence to conclude that the proportion of customers who selected products and then cancelled their transaction is less than 0.50 with the new system.

1. Let  be the population proportion of customers who select products and then cancel their transaction

H0: 

H1: 

n =100> 30, np=42>5, n(1-p)=58>5

Sampling distribution of p is approximately normal

 Z-test (Lower Tail)

With, 

Reject H0 if Z < −2.33



Since Z =-1.6 > -2.33

Do not Reject H0 at 

There is insufficient evidence to conclude that the proportion of customers who selected products and then cancelled their transaction is less than 0.50 with the new system.

1. The larger the sample size, the smaller is the standard error. Even though the sample proportion is the same value at 0.42 in (a) and (b), the test statistic is more negative while the p-value is smaller in (a) compared to (b) because of the larger sample size in (a).

**Q8**

a)

(i) Let π be the population proportion of stores carried the brand



 np=11.9>5 n(1-p)=73.1>5

 Sampling distribution of p is approximately normal.

 Critical Value 

Reject if 





We do not reject. There is insufficient evidence that Grant has poorer distribution in Mainland China than it does in Hong Kong.

(ii) There is insufficient evidence that Grant has poorer distribution in Mainland China than it does in Hong Kong.

(iii) p-value==0.1190

Reject  if p-value <0.05

 p-value = 0.1190> 0.05

 We do not reject  and making the same decision as in (i)

(iv) Choose 

For a fixed sample size, a larger value of would correspond to a smaller value of, that can decrease the penalty of committing type II error.

b)

(i) Standard error of sample proportion==0.0285

(ii)  =3000.42126 = 

 Sampling distribution of is approximately normal.

 = 

= = 0.6111

(iii) The range of 0.41 to 0.43 is more likely to lie because this range contains the population proportion that is 0.42.

(iv) In the sample size 300, 61.11% of sample will be expected to have the sample proportions between 0.4 and 0.45.

**Q9**

a) Let  be the proportion of unemployment rate of Hong Kong in 2002





sampling distribution p is normal

assume population follows binomial.

For 95% Confidence Interval,



 We are 95% confident that the population proportion of Hong Kong unemployment rate is estimated to be between 0.0672 and 0.0782.

b) Sample size

=64750

c) Letbe the proportion of unemployment rate of Shatin in 2002



=620>30,

np = 34 >5

n(1-p) = 586 >5

 the sampling distribution of is approximately normal



~

Reject if < -1.645

 == -1.738<-1.645

Therefore we reject. There is sufficient evidence that the unemployment in Shatin is lower than Hong Kong in 2002.

**Q10**

a)

(i) = 0.25

(ii) Possible value of sample proportion of preferring the brand:



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Not prefer | Not prefer | Not prefer | prefer |
| Not prefer | 0 | 0 | 0 | 0.5 |
| Not prefer | 0 | 0 | 0 | 0.5 |
| Not prefer | 0 | 0 | 0 | 0.5 |
| prefer | 0.5 | 0.5 | 0.5 | 1 |

Probability distribution of the sample proportion:

|  |  |  |  |
| --- | --- | --- | --- |
| p | 0 | 0.5 | 1 |
| Pr(p) |  |  |  |

Possible value of p from the above table: 0, 0.5, 1

(iii) p =0\*0.5625+0.5\*0.375+1\*0.0625=0.25

From part (i), =0.25

=> p =

=> sample proportion p is an unbiased estimator for 

(iv) n=2 < 30, np=2\*0.25=0.5<5, n(1-p)= 2\*(1-0.25)=1.5<5

=> sample distribution of p does not follow a normal distribution

b)

(i) H0:  5% vs H1::>5%

Type I error: let unhealthy patient with 5% white blood cells leave, resulting in unhealthy patients are not treated

Type II error: refer healthy patient with > 5% white blood cells to doctor, resulting in more cost is incurred or more patients are sent to doctor

Comparing the above type I and II error, type I error is more serious. We would rather make a type II error.

(ii) If there is no information available from past data,

α=1-90%=0.10,  = 1.645

E=1.95%=0.0195

0.5

Sample size, n==1779.11 (round up)

**Q11**

X is the number of passengers responded to the survey

π is the population proportion of passenger responded to the survey

a)  (round-up)

b) H0: π ≥ 0.13

H1: π < 0.13

As n = 350 > 30; np = 28 > 5; n(1-p) = 322 > 5

* p ~ N
* use Z test

At α = 0.05, reject H0 if Z < -1.645



Z = 

As Z = -2.7815 < -1.645, reject H0.

There is sufficient evidence that the response rate has been dropped.

c) 95% CI for π

= === [0.0516, 0.1084]

We are 95% confident that the true unknown population proportion of passengers who are willing to response to the survey is between 0.0516 and 0.1084 (i.e. 5.16% or 10.84%).