**The following are the temperature of effluent (discharge from a sewage treatment facility) on 24 consecutive days. Draw a box plot and interpret the result.**

**43, 47, 51, 48, 52, 50, 46, 49, 45, 52, 46, 51, 44, 49, 46, 51, 49, 45, 44, 50, 48, 50, 49, 50**

Arranging data,

Minimum Value = 43

Q1 = when i = 1, Q1 = 6th observation = 46

Q2 = when i = 2, Q2 = 12th observation = 49

Q3 = when i = 3, Q3 = 18th observation = 50

Maximum Value = 52, IQR = Q3 – Q1 = 4

Q3+ 1.5 (IQR) = 50+1.5\*4 =56

Q1- 1.5 (IQR) = 46-1.5\*4 = 40

UPPER whisker = 52

Lower whisker = 43

Since, upper whisker is equal to max. value and lower whisker is equal to minimum value, so there are no outliers.

DRAW BOX PLOT:

**A civil engineer monitors water quality by measuring the amount of suspended solids in a sample of river water. Over 11 weekdays, he observed 14, 12, 21, 28, 30, 63, 29, 65, 55, 19, 20 suspended solids (parts per million). Find the third quartile and interpret its meaning**

ARRANGE DATA:

Q3 = when i = 3, Q3 = 8.25TH = 8th observation + 0.25\*(9th – 8th obs.) = 36.25

Qi =

Di =

Pi =

CV of height = 4.41%

CV of weight = 15.25%

Weight is more variable than height.

**The mean and standard deviation of 20 items is found to be 10 and 2 respectively. At the time of checking it was found that one item 8 was incorrect. Calculate the mean and standard deviation if : a) the wrong item is omitted b) it is replaced by 12.**

No. of observations, N = 20

Mean, = 10

Standard deviation, σ = 2

Wrong observation = 8

Mean =

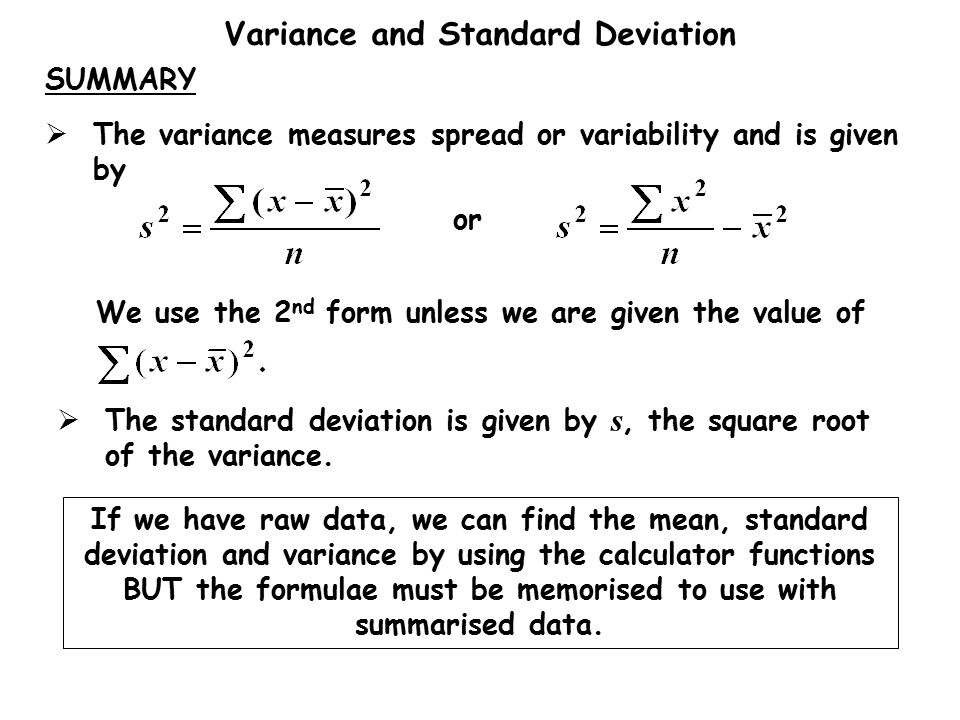
= N \* = 20 \* 10 = 200

If wrong observation 8 is omitted,

Correct = 200 – 8 =192

Correct mean = = = 10.105

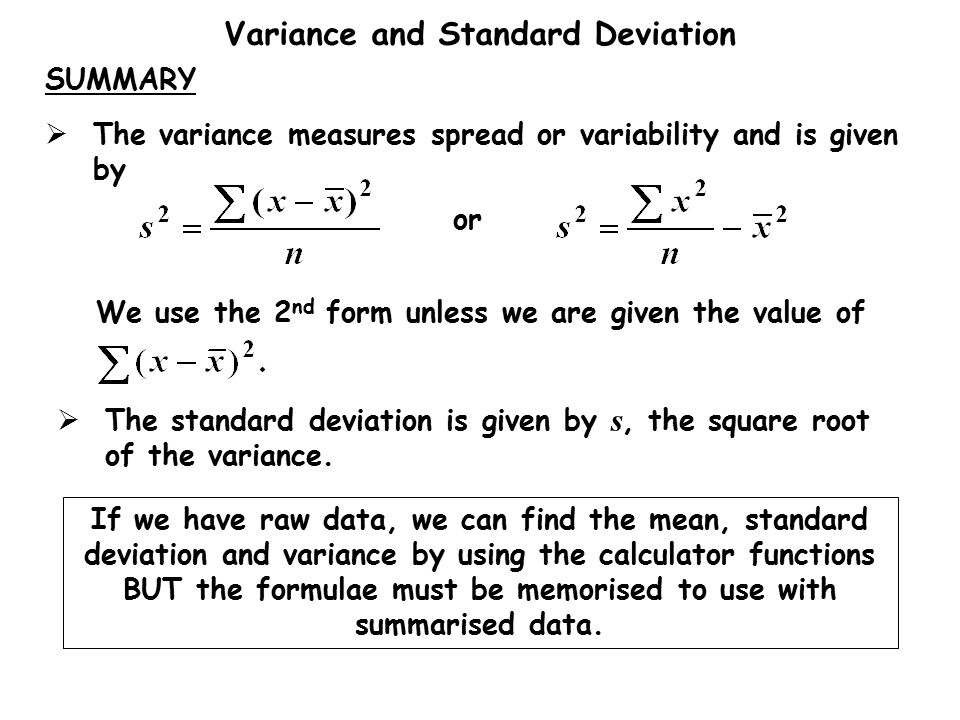
For standard deviation,



22 = - 102

So, = 2080

After omitting wrong observation, Correct = 2080 - 82 = 2016

Now correct σ2  = - (10.105)2

σ = = 2.023

=

If C is added to each observation,

= = = + c

Combined mean () =

Prob (2 red cards) = =

**From the following data state which is more consistent and why?**

|  |  |  |
| --- | --- | --- |
|  | Sample 1 | Sample 2 |
| Age | 25 years | 11 years |
| Mean Weight | 145 pound | 80 pound |
| Standard deviation | 10 pound | 10 pound |

CV for sample 1 = 10/145\*100% = 6.896%

CV for sample 2 = 10/80\*100% = 12.5%

Since, CV of sample 1 < CV of sample 2, sample 1 is more consistent.

* A committee of 6 members is to be formed out of a group consisting 7 men and 4 women. Calculate the probability that committee will consist of (a) exactly 2 women and (b) at least 2 women.

Total = 11 people (7men and 4 women)

We have to select 6 members.

1. P(2 women, 4 men) =
2. P(at least 2 women) = P(2women, 4 men) + P(3women, 3 men) + P(4 women, 2 men) = + + = ……..

* **Five people are being considered for three awards, and no person can receive more than one award.**

1. **In how many ways can these awards be given? [60]**
2. **If three of these people are city officials, in how many ways could the awards be given to these officials? [6]**
3. **If all candidates are equally qualified for the three awards, what is the probability that the awards will be presented to the three city officials. [1/10]**
4. P(n,r)= P(5, 3) = 60
5. 3\*2\*1 = 6= P(3,3)
6. P(city officials will get awards) = 6/60= 1/10

In a small town, population was categorized as follow

|  |  |  |
| --- | --- | --- |
|  | **Employed (E)** | **Unemployed (U)** |
| **Male (M)** | **460** | **40** |
| **Female (F)** | **140** | **260** |
| **Total** | **600** | **300** |

One person is selected at random.

1. If the person is found to be employed, what is the probability that the person was male? P(M/E) = 460/600 =
2. If the person was female, what is the probability that she was unemployed?

P(U/F)= 260/400

1. What is the probability that the person was employed? = P(E)= 600/900=6/9
2. What is the probability that the person was male? =

= P(M) = P(M∩E)+P(M∩U) = 460/900+40/900= 500/900= 5/9

**Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products. What is the probability that a product attains a good review?**

Solution:

E1 = event of selecting highly successful product

E2 = event of selecting moderately successful product

E3 = event of selecting poor product

A = event of **product receiving good review**

P(E1) = Probability of selecting highly successful product = 0.4

P(E2) = Probability of selecting moderately successful product = 0.35

P(E3) = Probability of selecting poor product= 0.25

P(A/E1) = Probability of receiving good review if the product is highly successful = 0.95

P(A/E2) = Probability of receiving good review if the product is moderately successful = 0.6

P(A/E3) = Probability of receiving good review if the product is poor = 0.1

P( A) = P(A and E1) + P(A and E2) + P(A and E3)= = 0.61

**There are three boxes having the following compositions of black and white balls:-**

**Box I – 7 white, 3 black, Box II – 4 white, 6 black, Box III – 2 white, 8 black.**

**One of these boxes is selected at random with probabilities 0.20, 0.60 and 0.20 respectively. From the selected box, two balls are drawn at random (without replacement). Calculate the prob. that both these balls are white. (0.177)**

E1……………………..

E2……………..

E3 ………………

A = event of selecting two white balls

P(E1)=0.2, P(E2) = 0.6, P(E3) = 0.2

P(A) = ? = P(A∩E1) + P(A∩E2 ) + P(A∩E3) =

P(A/E1) = prob. Of selecting two white balls from box I = =

P(A/E2)=……

P(A/E3) =……

Tutorial 1

1. Let Missing frequencies be x and y

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Expenditure | 40-59 | 60-79 | 80-99 | 100-119 | 120-139 |
| No. of families | 50 | X | 500 | y | 50 |
| Cumulative freq. | 50 | 50+x | 550+x | 550+x+y | 600+x+y |

Total family = 1000

50+x+500+y+50 = 1000

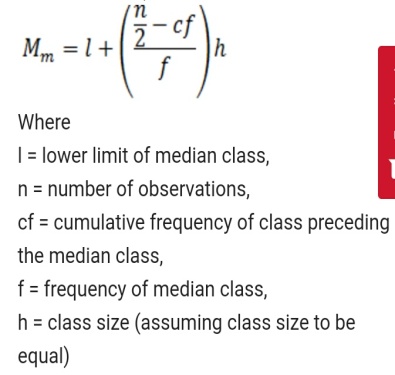
600+x+y = 1000

X+ Y = 400…………………………..eq (1)

**If the median of frequency distribution is 87,**

**It lies in 80-99 class.**

**So, l = 80, f = 500, c.f.= 50+x**

**Median = **

**Solving, x = 275 So, y = 400-275= 125**

**Q. 3. Compare Coefficient of variation of two micrometers.**

**Whichever is small, it is precise.**

**Q4. ii) one is good, one is defective = (A∩Bc) ∪ (Ac∩B)**

**P(A) = 0.8, P(B/A) = 0.85, P(B/Ac) = 0.75**

a) The second component is good = P(B) = P(A∩B) + P(Ac ∩ B) = P(B/A)\*P(A) + P(B/Ac)\*P(Ac) = 0.83

b) At least one component is good = P(A∪ B) = P(A) + P(B) – P(A∩B) = 0.95

c) Are they independent? Verify your answer. [if A and B are independent, then

P(A∩B) = 0.68

P(A)\*P(B) = 0.664

Since P(A∩B) and P(A)\*P(B) are not equal, so A and B are not independent.

**The probability that a person at a service station will get tire checking is 0.12, the probability that he will get oil check is 0.29 and the probability that he will get both checking is 0.07**

1. **a. What is probability that a person will have neither tire nor oil checking? [0.66]**
2. **b. Find the probability a person will get tire checking given he has his oil checking?**

T = event of tire checking

O = event of oil checking

P(T)= 0.12, P(O) = 0.29, P(T∩O) = 0.07

1. P(T∪O)c = 0.66
2. P(T/O) =

Let Pi = 35, find the value of “i”

Pi = *l* +

35 = 30 +

Solving,

i = 25

Pass% of students = 100-25 = 75%

For 34, i = 23.25

For 79, i = 91.625

3 economist, 4 engineers, 2 statisticians and 1 doctor

What is the probability that the committee consist of the doctor and at least one economist?

Possible cases are

(i) 1 doctor, 1 economist, 2 from Engineers and Statisticians ⇒ P(i) = 1C1\*3C1\*6C2/10C4

(ii) 1 doctor, 2 economists, 1 from Engineers and Statisticians⇒ P(ii) = 1C1\*3C2\*6C1/10C4

(iii) 1 doctor, 3 economists, 0 from Engineers and Statisticians⇒ P(iii) = 1C1\*3C3/10C4

Required answer = ADD

Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products. What is the probability that a product attains a good review?

E1 = event of selecting highly successful product

E2 = event of selecting moderately successful product

E3 = event of selecting poor product

A = product receiving a good review

P(E1)= prob. Of selecting highly successful product = 0.4,

P(E2) = 0.35 and

P(E3) = 0.25

P(A/E1)= prob. Of getting good review if it is highly successful product = 0.95

P(A/E2) = prob. Of getting good review if it is moderately successful product = 0.6

P(A/E3) = prob. Of getting good review if it is poor product = 0.1

P(A) = P(A∩E1) + P(A ∩E2) + P(A∩E3) = =

Three different airlines (1,2,3) operate night flights from Kathmandu to Pokhara. Experience has shown that 40% of airline 1 are late in takeoff, 50% of airline 2 are late in take off and 70% of airline 3 flights are late in take off. On a particular night, a person selects randomly one of the three airlines and flies on its night flight. **a) What is the prob. that he selects airline 1 and he will be late in take off?** b) What is the prob. that he will be late in taking off? c) If he was late in take off, what is the prob. that he selected air line 1? (0.133, 0.533, 0.25)

P(E1) = prob. of selecting airlines 1 = 0.333 =1/3

P(E2) = 1/3

P(E3) = 1/3

A = event of late take off

P(A/E1) = 0.4 ………………….

P(A/E2) = 0.5 and P(A/E3) = 0.7

1. P(E1∩A) = P(A/E1)\*P(E1)
2. P(A) = ? P(E1∩A)+ P(E2∩A) + P(E3∩A)
3. P(E1/A) = …….formula……

There are three boxes having the following compositions of black and white balls:- Box I – 7 white, 3 black, Box II – 4 white, 6 black, Box III – 2 white, 8 black. One of these box is selected at random with probabilities 0.20, 0.60 and 0.20 respectively. From the selected box, two balls are drawn at random (without replacement). Calculate the prob. that both these balls are white.

Mean = = + + ……..

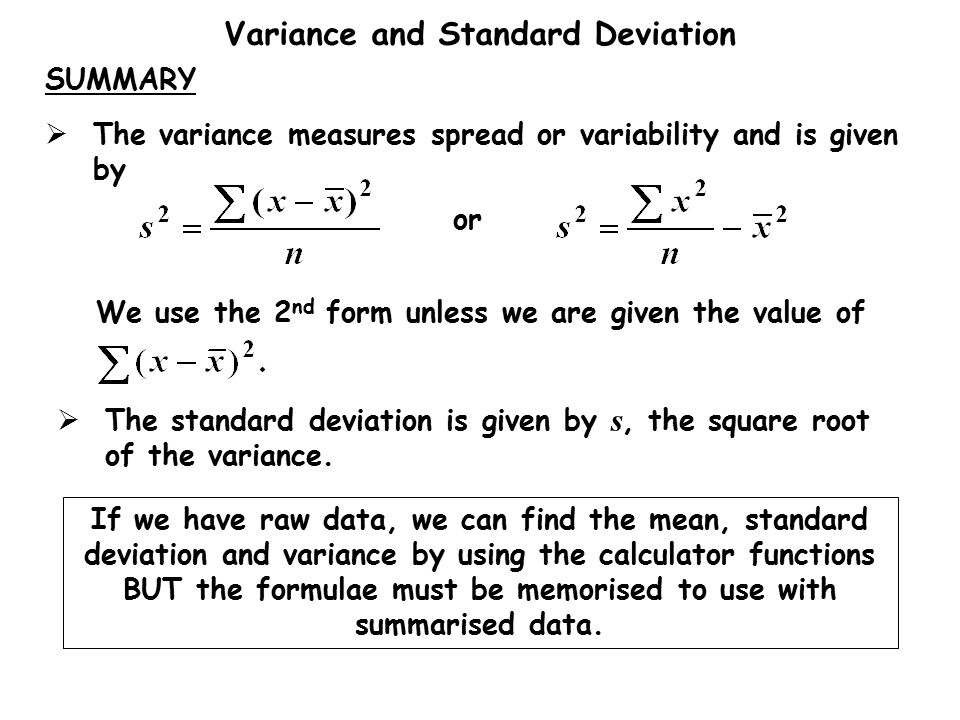
Expected value of Random Variable

E(X) =

**In 3 tosses of coin, no. of possible outcomes = {TTT, HTT, THT, TTH, HHT, HTH, THH, HHH}**

|  |  |
| --- | --- |
| X (no. of heads) | P(X) |
| 0 | 1/8 |
| 1 | 3/8 |
| 2 | 3/8 |
| 3 | 1/8 |
| Total | 1 |

EXPECTED NO. OF HEADS E(X) = 0\*1/8 + 1\*3/8 + 2\*1/8 + 3\* 1/8 = 3/2 = 1.5

σ 2 = - Mean2

= + +………… - Mean2 = - [E(X)]2

= E(X2) – [E(X)]2

P(X≥ 2) = P(X=2)+P(X=3)+……..+P(X=7) = 1 – P(X ≤ 1)

Expected gain = 0.3\* 4000 – 0.7\*1000 = 500

Expected gain is zero, i.e. neither loss nor profit

From a lot of 10 items containing 3 defectives, a sample of 4 items is drawn at random. Let the random variable X denote the no. of defective items in the sample. Answer the following when the sample is drawn (without replacement)

**Find the probability distribution of X**

Find P(X ≤ 1), P(X < 1) and P(0 < X < 2)

Obtain the mean and variance for the distribution.

Let r.v. X = no. of defective items in sample of 4

= 0, 1, 2, 3

P(X=0)= prob. Of no defectives in sample = =

P(X=1)= prob. Of 1 defective in sample = =

P(X=2)= prob. Of 2 defectives in sample = 3/10

P(X=3) = prob. Of 3 defectives in sample= 1/30

The prob. Distribution is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | 0 | 1 | 2 | 3 |
| P(X) | 1/6 | 1/2 | 3/10 | 1/30 |

Find the probability distribution of boys in the family of 3 children assuming equal probability for boys and girls.

Let RV X = Number of boys in the family of 3 children = 0, 1, 2, 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RV X (No. of boys) | 0 | 1 | 2 | 3 |
| Probability, P(X) | 1/8 | 3/8 | 3/8 | 1/8 |

P(B) = 1/2 = P(G)

P(BBB) = P(B)\*P(B)\*P(B)= ½\*1/2\*1/2 = 1/8

P(BBG) + P(BGB) + P(GBB) =

BGG, ……………………………..

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x: | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| P(x): | a | 3a | 5a | 7a | 9a | 11a | 13a | 15a | 17a |
| Cumulative Prob.  P(X≤ x) | a | 4a | 9a | 16a | 25a | 36a | 49a | 64a | 81a |

**5 tosses : 3 heads and 2 tails appear**

1. P(H T H H T) = P(H)\*P(T)\*……\*P(T) = p\*q\*…q= p3q2
2. P(H H T T H) = p3q2
3. P(H T T H H) = p3q2
4. ……..
5. …….

Total = C(n,x) pxqn-x

In C(n,x) ways, x success can occur out of n trials.

Expected value of Binomial Random Variable

E(X) = = = ………………………= np

Find the probability distribution of boys in the family of 3 children assuming equal probability for boys and girls

Let RV X = No. of boys in family of 3 children = 0, 1, 2, 3

|  |  |
| --- | --- |
| X (no. of boys) | P(X) |
| 0 | 1/8 |
| 1 | 3/8 |
| 2 | 3/8 |
| 3 | 1/8 |
| Total | 1 |

**In family 3 children, no. of possible outcomes = {GGG, BGG, GBG, GGB, BBG, BGB, GBB, BBB}**

P(X=0)= P(GGG) = P(G)\*P(G)\*P(G) = ½\*1/2\*1/2= 1/8

From a lot of 10 items containing 3 defectives, a sample of 4 items is drawn at random. Let the random variable X denote the no. of defective items in the sample. Answer the following when the sample is drawn (without replacement)

Find the probability distribution of X

Total = 10, defectives = 3, Non defectives = 7

Sample (n) = 4

RV X = No. of defective items = 0, 1, 2, 3

P(X =0) = prob. Of all good ones = 7C4 / 10C4 = 1/6

P(X= 1) = prob. Of 1 defective and 3 non defectives = 3C1\*7C3/10C4 = ½

P(X=2) = prob. Of 2 defectives and 2 non defective = 3C2\*7C2 / 10C4 = 3/10

P(X=3) = prob. Of 3 defectives and 1 non defective = 3C3\*7C1/10C4 = 1/30

|  |  |
| --- | --- |
| X | P(X=x) |
| 0 | 1/6 |
| 1 | ½ |
| 2 | 3/10 |
| 3 | 1/30 |
| Total | 1 |

E(X ) = Mean = Σ X\* P(X=x) = 1.2

E(X2)= Σ X2\* P(X=x) = 2, now Variance = E(X2) – [E(X)]2 = 2- 1.44= 0.56

**In a family of 8 children, what is the probability of getting exactly 3 girls?**

Let R.V. X = No. of girls in family of 8 children, **which is a binomial random variable**

Number of trials (n) = 8

Prob. Of being girl (p) = 0.5,

Prob. Of being boy (q) = 1-p = 0.5

P(X = 3) = ?

Using Binomial prob. Distribution,

P(X=x) = C(n,x)pxqn-x

P(X=3) = C(8,3)(0.5)3(0.5)8-3 = [use calculator] =

A student selects his answer on true/false exam by tossing a coin (so that any particular answer has a 0.5 prob. of being correct). He must answer at least 70% correctly in order to pass. Find the probability of passing when the number of questions is 10.

Random variable X = No. of correct answers, **which is a binomial r.v.**

No. of trail, n = 10

Prob. Of getting correct answer, p = ½

Prob. Of getting wrong answer, q = ½

P(X ≥ 7) = P(X=7) + P(X=8) + P(X=9) + P(X = 10)

= C(10, 7) (1/2)7 (1/2)10-7 + …………………….= use calculator = 0.172

= = use calculator

A multiple-choice test consists of 15 questions and for each question; there are 5 possible answers. If for each question, an answer is selected in a completely random fashion, what is the probability that (a) exactly 8 (b) at least 2 are answered correctly?

Random variable, X = No. of correct answers, **which is a binomial random variable**

Prob. Of getting correct answer, p = 1/5

Prob. Of getting wrong answer, q = 4/5

No. of questions, n= 15

1. P(X=8) = C(15, 8) (1/5)8 (4/5)15-8
2. P(X ≥ 2) = P(X=2) + …………………..+ P(X=15) = 1 –P(X ≤ 1) = 1 – [P(X= 1) + P(X= 0)]

Average no. of correct answers = E(X) = np =

Out of 9000 families 4 children each, how many families would you expect to have (i) 2 boys and 2 girls (ii) at least one boy (iii) no girls (iv) at most 2 girls (v) at least one boy and one girl.

Random variable, X = no. of boys, which is a binomial r.v.

No. of trials, n = 4,

Prob. Of getting a boy, p = ½

Prob. Of getting a girl, q = ½

1. P(X= 2) = 3/8

No. of families in which there are 2 boys and 2 girls= 3/8 \* 9000 = 3375

1. P(X ≥ 1) = 1- P(X = 0)
2. P(X = 4)
3. At most 2 girls

P(X = 4) + P(X= 3) + P(X= 2)

1. At least one boy and one girl = P(X=1) + P(X=2) + P(X=3)

**NEGATIVE BINOMIAL DISTRIBUTION**

**r = 2 heads**

H H

T H H

H T H

H T T **H**

T H T **H**

T T H **H**

Required success = “r”

**S S F F S S F…….…….F F S**

**P(S S F F S S F…….…….F F S)**

P(S)\*P(S)\* P(F)\*……………P(F)\*P(F) \* P(S)

X failures and (r-1) success rth success

**C(x+r-1, r-1) = C(x+r-1, x)**

Random variable X= no. of failures before rth success is obtained

r = no. of success which is fixed

x+r = no. of trials, which is random

A scientist inoculates several mice, one at a time, with a disease germ until he finds two, which have contracted the disease. If the prob. of contracting the disease is 1/6, what is the prob. that eight mice are required?

**O O O O O O O O**

Let r.v X = no. of mice which do not have disease, **which is a Negative binomial random variable**.

Required no. of success, r = 2

Prob. of contracting disease, p = 1/6

Prob. of not contracting disease, q = 5/6

P(X +r = 8) = P(X = 6) = C(X+r-1, r-1)prqx =

Three people toss a coin and odd man pays for the coffee. If the coins all turn up the same, they toss again. Find the prob. that fewer than 4 tosses are needed.

Let R.V . X = no. of tosses in which same face turns up, which is a Negative Binomial r.v.

Prob. of success , p = 6/8

Prob. failure, q = 2/8

Required no. of success, r = 1

P(X +r < 4) = ???

= P(X < 3) = P(X=0) + P(X=1) + P(X=2) = C(x+r-1,r-1) prqx

**In 3 tosses of coin, no. of possible outcomes = {TTT, HTT, THT, TTH, HHT, HTH, THH, HHH}**

A library has 20 copies of textbooks of which 8 are 1st edition and 12 are 2nd edition. The teacher has requested 5 copies of that book. If the books are selected in random, what is the probability that 2 of those selected are 2nd edition?

RV X= no. of 2nd edition books in sample n, which is a Hypergeometric rv

Total number of books, N= 20

No. of 2nd edition books, M= 12

No. of 1st edition books, N-M = 8

Required no. of books, n = 5

No. of 1st edition books in sample, n-x = 3

P(X = 2) = ?

Among 300 employees, 240 are union members, while the others are not. If 8 of the employees are chosen by lot to serve on the committee, find the probability that 5 of them will be union members while the others are not using

a) Hypergeometric distribution b) Binomial distribution.

N = 300, M = 240, N-M = 60

n = 8,

P(X= 5) = C(240,5) \* C(60,3) / C(300, 8)

Since, n/N = 8/300 = 0.0267 < 0.1 we can use Binomial distribution also

P(X= 5) = C(8,5) (240/300)5 (60/300)8-5

E(X+r) = E(X+1) = E(X) + 1 = r(1-p)/p + 1

A library has 20 copies of textbooks of which 8 are 1st edition and 12 are 2nd edition. The teacher has requested 5 copies of that book. If the books are selected in random, what is the probability that 2 of those selected are 2nd edition?

Random Variable, X = No. of 2nd edition books in sample of 5, **which is a Hypergeometric r.v.**

Total number of books, N= 20

No. of 2nd edition books, M= 12

No. of 1st edition books, N-M = 8

Sample, n = 5

P(X = 2) = ?

X = 0, 1, 2, 3, 4, 5

P(X=x ) =

Among 300 employees, 240 are union members, while the others are not. If 8 of the employees are chosen by lot to serve on the committee, find the probability that 5 of them will be union members while the others are not using a) Hyper geometric distribution b) Binomial distribution

N = 300

M = 240

N-M = 60

n = 8

P(X= 5) = ?

n/N = 8/300 = 0.0266 < 0.1we can use Binomial distribution as an approximation to Hypergeometric prob. distribution.

Using binomial,

n = number of employees chosen for committee = 8

prob. of success, p = 240/300 = 4/5

prob. of failure, q = 1/5

P(X = 5) = C(8, 5) (4/5)5(1/5)8-5

Suppose we are investigating the safety of a dangerous intersection of a road. Past police records indicate a mean of 5 accidents per month at this intersection. Suppose the number of accidents is distributed according to a Poisson distribution. Calculate the probability in any month of exactly 0 and between 2 to 4 accidents.

Let R.V X = no. of accidents occurring per month, which is a Poisson distribution.

λ = average number of accidents in the intersection = 5

P(X=0) = =

P(2 ≤ X ≤ 4) =

A manufacturer of matchstick knows that on average 2% of his production is defective. He sells matchsticks in boxes of 100 and guarantees that not more than 2 matchsticks will be defective. What is the probability that a match box selected at random will meet the guaranteed quality?

r.v. X = no. of defective matchsticks in a box, which is a Poisson r.v.

Prob. of matchstick being defective, p = 0.02

Number of matchsticks in a box, n = 100

Average number of defective matchstick, λ = 100 \* 0.02 = 2

Now, P( X ≤ 2) = P(X=0) + P(X=1) + P(X=2)

In a certain factory, there is 0.2% probability for any blade to be defective. Blades are supplied in packets of 10. Using Poisson distribution, calculate the approximate no. of packets containing (a) no defective (b) one defective (c) two defective respectively in a consignment of 20,000 packets.

Number of blades in packet, n = 10

Prob. of defective blades, p = 0.002

Avg. number of defective blades, λ = 10\*0.002 = 0.02

1. P(X=0) \* 20,000

At a checkout counter, customers arrive at an average rate of 1.5 per minute. Find the probability. that (a) at most four will arrive in any given minute. (b) at least three will arrive during an interval of 2 minutes (c) at most 15 will arrive during an interval of 6 minutes.

r.v. X = number of customers arriving at checkout counter, which is a Poisson r.v.

1. λ = 1.5, P(X ≤ 4)
2. λ = 1.5 \*2=3,

P(X ≥ 3) = 1 – P(X<3)

1. λ = 1.5\*6= 9 , P(X ≤ 15) = =

The probability of a serious fire during a given year to any one house in a particular city is believed to be 0.005. A particular insurance company holds fire insurance policies on 1000 homes in this city.

(a) Find the probability that the company will not have any serious fire damage claims by the owners of these homes during the next year

(b) Find the probability they will have no more than three claims.

f(x) = Ce-3x

Find C

= 1

C= 3

P(1 < X < 2) = = 0.047

If the city’s power plant has a daily capacity of 12 million kilowatt-hours, what is the probability that this power supply will be inadequate on any given day?

P(X ≥ 12) = =

= 1 - P(X < 12) =

r.v. X = the time which elapses between the bell and the end of the lecture

P(X≤ ½) =

P(0.25≤ X ≤ 0.5)

P(X ≥ 2/3) =

A player tosses 3 fair coins. He wins Rs.10, if 3 heads appears, Rs.6 if 2 heads appear and Rs.2 if one head appears. On the other hand, he loses Rs.25, if 3 tails appear. Find the expected gain of the player.

r.v. X = gain of the player

P(3 heads) = 1/8

P(2 heads) = 3/8

P(1 head) = 3/8

P(0 head) = 1/8

Expected gain = Rs.10\* + Rs 6\* + Rs 2\* - Rs 25\* = Rs. 1.125

A multiple-choice test consists of 15 questions and for each question; there are 5 possible answers. If for each question, an answer is selected in a completely random fashion, what is the probability that (a) exactly 8 (b) at least 2 are answered correctly?

Random variable, X = number of correct answers, **which is a binomial random variable**

Number of questions, n = 15

Prob. of getting correct answer, p= 1/5

Prob. of getting wrong answer, q = 4/5

1. P(X =8)= C(15, 8) (1/5)8 (4/5)15-8
2. P(X ≥ 2) = 1 – P(X < 2) = 1 – P(X=0) – P(X=1)

Out of 9000 families 4 children each, how many families would you expect to have (i) 2 boys and 2 girls (ii) at least one boy (iii) no girls (iv) at most 2 girls (v) at least one boy and one girl

Random variable, X = number of boys in a family of four, **which is a binomial random variable**

Number of trials, n = 4

Prob. of getting a boy, p = ½

Prob. of getting a girl, q = ½

1. P(X=2)= prob. of getting two boys = C(4, 2) (0.5)2(0.5)4-2 = 3/8

No. of families having two boys = 9000\*3/8 = 3375

1. P(X ≥ 1) \*9000
2. P(X=4) \*9000
3. [P(X=4)+ P(X=3) + P(X=2)] \* 9000
4. [P(1≤ X ≤ 3) ] \* 9000

Suppose we are investigating the safety of a dangerous intersection of a road. Past police records indicate a mean of 5 accidents per month at this intersection. Suppose the number of accidents is distributed according to a Poisson distribution. Calculate the probability in any month of exactly 0 and between 2 to 4 accidents

Random variable, X = number of accidents per month, which a Poisson R.V

λ = average number of accidents per month = 5

P(X=0) = =

P(2 ≤ X ≤ 4) =

A manufacturer of matchstick knows that on average 2% of his production is defective. He sells matchsticks in boxes of 100 and guarantees that not more than 2 matchsticks will be defective. what is the probability that a match box selected at random will meet the guaranteed quality?

Random variable, X = number of defective matchsticks, which is a poisson r.v.

Prob. of a matchstick being defective, p = 0.02

Number of matchsticks in a box, n = 100

λ = avg. no. of defective match = np= 100\*0.02= 2

P(X ≤ 2) = P(X=0) + P(X=1) + P(X=2)

In a certain factory, there is 0.2% probability for any blade to be defective. Blades are supplied in packets of 10. Using Poisson distribution, calculate the approximate no. of packets containing (a) no defective (b) one defective (c) two defective respectively in a consignment of 20,000 packets.

r.v. X = number of defective blades in a pack, which **is Poisson r.v.**

Prob. of blade being defective, p = 0.002

Number of blades in a packet, n= 10

λ = np = 0.02

1. P(X=0) \* 20,000 =
2. P(X= 1) \*20,000 =
3. P(X = 2) \*20,000 =

At a checkout counter, customers arrive at an average rate of 1.5 per minute. Find the probability. that (a) at most four will arrive in any given minute. (b) at least three will arrive during an interval of 2 minutes (c) at most 15 will arrive during an interval of 6 minutes

1. λ = 1.5 customers per minute, P(X≤ 4) = ??
2. λ = 2\*1.5 = 3 customers per minute, P(X ≥ 3) = 1 – P(X ≤ 2)
3. λ = 6\*1.5 = 9 customers per minute, P(X ≤ 15) = =

A fair dice was rolled until one gets a Six. Find the expected number of toss required.

Required success, r = 1

R.v.X = number of toss not getting six, which is a negative binomial r.v.

Prob. of of getting success, p = 1/6

Number of toss = X + r

E(X+r) = E(X+1) = E(X) + 1 = r(1-p)/p + 1 = 6

A scientist inoculates several mice, one at a time, with a disease germ until he finds two, which have contracted the disease. If the prob. of contracting the disease is 1/6, what is the prob. that eight mice are required?

r.v. X = number of mice without disease germ, which is a negative binomial r.v.

Required success, r = 2

Prob. of getting mice that have disease, p = 1/6

P(X+r = 8) = C(X+r-1, r-1)prqx  =

O **O** O O O O O **O**

Three people toss a coin and odd man pays for the coffee. If the coins all turn up the same, they toss again. Find the prob. that fewer than 4 tosses are needed.

R.V. X = number of times coins turning up same, which is a negative binomial r.v.

Required success, r = 1

Prob. of getting odd results, p = 6/8

Prob. of getting no decision, q = 2/8

P(X+r < 4) = P(X < 3) = P(X=0) + P(X=1) +P(X=2) = =

**In 3 tosses of coin, no. of possible outcomes = {TTT, HTT, THT, TTH, HHT, HTH, THH, HHH}**

Tutorial 2

Q.no. 5 Consider a group of five potential blood donors - A, B, C, D, and E - of whom only A and B have O+ type blood. Five blood samples, one from each individual, will be typed in random order until an O+ individual is identified. If the random variable X is number of typings necessary to identify an O+ individual, obtain the probability mass function of X.

A, B, C, D , E

RV X = 1, 2, 3, 4

P(X=1) = Prob. that A or B is selected in first typing = 2/5

P(X=2) = Prob. of selecting C,D or E first and then A or B = 3/5 \* 2/4

P(X=3) = Prob. of selecting C, D or E first and second, and then A, B) = 3/5\*2/4\*2/3

P(X=4) = Prob. of selecting C, D or E first, second and third and then A or B = 3/5\*2/4\*1/3\*2/2

The required prob. distribution is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RV X | 1 | 2 | 3 | 4 |
| P(X=x) | 2/5 | 6/20 | 12/60 | 12/120 |

Five coins are tossed 320 times. If coins are unbiased, construct the probability distribution table of number of heads obtained. Also find the mean, variance of the probability distribution.

RV X = no. of heads in 5 tosses = 0, 1, 2, 3, 4, 5

Since coins are unbiased, prob. of getting head, p = 0.5, q = 0.5, n = 5

RV X ~ Binomial distribution

P(X=0) =

P(X=1) =

P(X=2) =

P(X=3) =

P(X= 4) =

P(X=5) =

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| RV X | 0 | 1 | 2 | 3 | 4 | 5 |
| P(X=x) | 1/32 | 5/32 | 5/16 | 5/16 | 5/32 | 1/32 |
| No. of heads in 320 tosses | 1/32\*320  =10 | 5/32\*320  =50 | 5/16\*320  =100 | 5/16\*320  =100 | 5/32\*320  =50 | 1/32\*320  =10 |

Mean = np

Variance = npq

**RV X= number of heads in 5 tosses, which is a Binomial r.v.**

**X = 0,1,2,3,4,5**

**Number of trials, n = 5**

**Prob. of getting head, p – 0.5, q= 0.5**

**P(X=x)= C(5,x)(0.5)x(0.5)5-x**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **X** | **0** | **1** | **2** | **3** | **4** | **5** |
| **P(X=x)** | **1/32** | **5/32** | **5/16** | **5/16** | **5/32** | **1/32** |
| **Number of heads in 320 tosses** | **320\*1/32** | **320\*5/32** | **320\*5/16** | **320\*5/16** | **320\*5/32** | **320\*1/32** |

**Mean, E(X) = n\*p= 5\*0.5 = 2.5**

A library has 20 copies of textbooks of which 8 are 1st edition and 12 are 2nd edition. The teacher has requested 5 copies of that book. If the books are selected in random, what is the probability that 2 of those selected are 2nd edition?

R. V. X = no. of 2nd edition books, which is a hypergeometric r.v.

Total number of books, N = 20

No. of 2nd edition books, M = 12

No. of 1st edition books, N-M = 8

Number books required, n = 5

P(X=2) = C(12, 2) \* C(8, 3) / C(20,5) = use calculator

Among 300 employees, 240 are union members, while the others are not. If 8 of the employees are chosen by lot to serve on the committee, find the probability that 5 of them will be union members while the others are not using a) Hyper geometric distribution b) Binomial distribution

R. V . X = number of union members in committee ~ hypergeometric distribution

N= 300

M = 240

N-M = 60

n = 8

P(X= 5) = using hypergeometric

We can use binomial prob. distribution also, because, n/N = 8/300 = 0.0266< 0.1

So, prob. of success, M/N = 240/300 = 4/5 = p

**Tutorial 3**

Given, f(x) = Ce-3x , x >0

To find C,

C = 3

P(1 < x < 2) = =

E(X) =

E(X2) =

V(X)

RV X = daily consumption of electric power

P(X > 12) = = 1 -

P(X>a) = 0.24

P(Z > ) = 0.24

P(Z > z1) = 0.24

P(Z ≤ z1) = 1-0.24= 0.76

From Z score table,

= 0.71

a = 4\*0.71+12 answer

The mean yield for one-acre plot is 662 kilos with a s.d. 32 kilos. Assuming normal distribution, how many one-acre plots in a batch of 1000 plots would you expect to have yield

1. over 700 kilos b) below 650 kilos

RV X= yield for one acre plot ~ N(μ, σ)

Mean yield, μ = 662 kg

Standard deviation, σ = 32 kg

1. P(X > 700) = P(Z > 1.19) = 1- P(Z ≤ 1.19) = 1 – 0.8830 = 0.117

So, number of 1 acre plots in which yield is over 700kg is 0.117\*1000 = 117

1. P(X < 650)

The time required to assemble a piece of machinery is a random variable having approximately a normal distribution with mean 12.9 minutes and standard deviation of 2 minutes. What are the probabilities that the assembly of a piece of machinery of this kind will take (i) at least 11.5 minutes (ii) between 11.0 to 14.8 minutes? [

RV X = time required to assemble machinery ~ N (μ, σ)

Mean , μ = 12.9

S.D, σ = 2

P(X ≥ 11.5)

13. 1 – P(5.9 ≤ X ≤ 6.1)

9. Of a large group of men, 5% are under 60 inches in height and 40% are between 60 and 65 inches. Assuming a normal distribution, find the mean height and standard deviation.

RV X = height of men ~ N (μ, σ)

According to question,

P(X < 60) = 0.05 ………………eq(1)

P(60 < X < 65) = 0.4 …………….eq (2)

From eq (1)

P(Z < ) = 0.05

From Z score table

= - 1.645 …………….eq (3)

From eq (2)

P( < Z < ) = 0.4

P(Z < ) – P(Z < ) = 0.4

P(Z < ) – 0.05 = 0.4

P(Z < ) = 0.45

From Z score table,

= -0.13 ………eq (4)

Solving eq (3) and (4)

We get

σ, μ

Let ‘a’ be the minimum pass marks

RV X = capacity of fuel tank

To travel the distance of 370 miles, the fuel tank should hold at least 370/25 = 14.8 gallons

P(X≥ 14.8) =??

Of a large group of men, 5% are under 60 inches in height and 40% are between 60 and 65 inches. Assuming a normal distribution, find the mean height and standard deviation.

Let r.v. X = height of men ~ N(μ, σ)

According to question,

P(X<60) = 0.05………………..eq (1)

P(60<X<65) = 0.4……………….eq (2)

From eq (1), we get

P(Z < ) = 0.05

From Z score table,

= -1.645 ……………………eq (3)

From eq (2)

P( < Z < ) = 0.4

F() – F() = 0.4 OR P(Z < ) – P(Z< ) = 0.4

P(Z < ) – 0.05 = 0.4

P(Z < ) = 0.45

From Z score table,

= -0.13 ………………….eq (4)

Solving equ (3) and (4) we get σ and μ

Let ‘a’ be the height beyond which 10% of students lie

P(X> a) = 0.1

P(Z >) = 0.1

P(Z < ) = 0.9

So, = 1.28

A---------------------------370 miles----------------------------B 1 gallon travel 25 miles

To travel 370 miles without refueling, the tank should hold at least 370/25 = 14.8 gallons of fuel.

P(X ≥ 14.8)= ????

A multiple-choice quiz has 80 questions each with four possible answers of which only one is correct answer. What is the prob. that sheer guesswork yields from 25 to 30 correct answers about which the student has no knowledge?

Let r.v. X = number of correct answers, which is binomial r.v.

Prob. of getting correct answer, p = ¼

Prob. of getting wrong answer, q = ¾

Number of question, n = 80

Here, np= 80\*1/4=20 > 10, nq= 80\*3/4 = 60>10

**Using Normal approximation to binomial distribution,**

P(25 ≤ X ≤ 30) = P(24.5< X < 30.5) [using Continuity correction]

= P( < Z < ) = P(1.16 < Z < 2.71) = ……

21. Assume that the number of asbestos particles in a squared meter of dust on a surface follows a Poisson distribution with a mean of 1000. If a squared meter of dust is analyzed, what is the probability that less than 950 particles are found?

P(X < 950) = ??

λ = 1000

Using Normal approximation to Poisson distribution,

P(X < 950) = P(Z< ) = P(Z< -1.58) = 0.0571

If a random variable X has the gamma distribution with α = 2 and β = 1, find P(1.8 ≤ X ≤ 2.4).

RV X ~ Gamma distribution (2, 1)

P(1.8 ≤ X ≤ 2.4) = =

Suppose that the reaction time X of a randomly selected individual to a certain stimulus has a standard gamma distribution with α = 2 seconds. What is the probability that the reaction time is more than 4 seconds?

RV X ~ Standard Gamma distribution

α = 2 seconds

P(X> 4)= ? xe-xdx

0.0228 \* 2000 = 46 approx.

P(X< 45)= 0.31

P(X > 64) =0.08

P(X<64) = 0.92

5. RV X ~ N(μ, σ)

Mean = 12, SD = 4

P(X≥ 20) = P(Z ≥ 2)= 1- P(Z < 2.00)= 1- 0.9772 answer

P(X ≤ 20) = P(Z ≤ 2) = 0.9772 answer

P(0 ≤ X ≤ 12) = P(-3 ≤ Z ≤ 0) = F(0) – F(-3) = P(Z≤0) – P(Z≤-3)

= 0.5 – 0.0013 answer

P(X> a) = 0.24

P(X ≤ a) = 1- 0.24 = 0.76

P(Z ≤ ) = 0.76

From Z score table

= 0.71

So, a = 4\*0.71 + 12

13. RV X = length of plastic rods ~ N(6, 0.02)

1. P(5.9 ≤ X ≤ 6.1)

Of a large group of men, 5% are under 60 inches in height and 40% are between 60 and 65 inches. Assuming a normal distribution, find the mean height and standard deviation

Let r.v. X = height of men in inches ~ N(μ, σ)

P(X < 60) = 0.05…………………….eq (1)

P(60 < X < 65) = 0.4 ……………….eq (2)

From eq (1)

P(Z < ) = 0.05

From Z score Table,

= - 1.645 …………………eq (3)

From eq (2)

P( < Z < ) = 0.4

P(Z < ) – P(Z < ) = 0.4

P(Z < ) – 0.05= 0.4

P(Z < ) = 0.45

From Z Score table,

= -0.13 …………………eq (4)

Solve eq (3) and (4)

In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution.

P(X < 45) = 0.31

P(X > 64) = 0.08

P(X <64) = 1-0.08

16. Let ‘a’ be the minimum pass marks

P(X ≥ a) = 550/674 =

A.........................370 miles…………………….B

RV X= fuel capacity of tank

To run 25 miles , fuel needed is 1 gallon

To run 370 miles, 370/25 gallon

To travel distance of 370 miles without refueling, the fuel tank should hold at least 14.8 gallons

P(X ≥ 14.8) =???

From a population of five members 3, 6, 9, 12, 15 draw all possible random sample of size 3 without replacement. Obtain the sampling distribution of sample mean, show that expectation of sample mean is equal to population mean and find the standard error of sample mean.

Population size, N= 5

{3, 6, 9, 12, 15}

Sample size, n= 3

Possible number of samples = C(5,3)= 10

The sampling distribution of mean is:

|  |  |  |
| --- | --- | --- |
| s.No. | Sample | Sample mean (ӯ) |
| 1 | 3,6,9 | 6 |
| 2 | 3,6,12 | 7 |
| 3 | 3,6,15 | 8 |
| 4 | 3,9,12 | 8 |
| 5 | 3,9,15 | 9 |
| 6 | 3,12,15 | 10 |
| 7 | 6,9,12 | 9 |
| 8 | 6,9,15 | 10 |
| 9 | 6,12,15 | 11 |
| 10 | 9,12,15 | 12 |

Now, population mean, μ = = 9

To show, E(ӯ) = μ

E(ӯ) = = 9

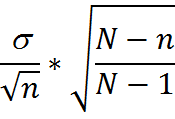
|  |  |  |
| --- | --- | --- |
| Sample mean (ӯ) | Freq. of means | Prob. distn of sample mean |
| 6 | 1 | 1/10 |
| 7 | 1 | 1/10 |
| 8 | 2 | 2/10 |
| 9 | 2 | 2/10 |
| 10 | 2 | 2/10 |
| 11 | 1 | 1/10 |
| 12 | 1 | 1/10 |
| Total | 10 | 1 |

E(ӯ) = (ӯ) = 6\*1/10+ ………………+ 12\*1/10 = 9 = μ

Standard error of sample mean, SE = ????

Here, n/N= 3/5 = 0.6 > 0.05

Population standard deviation, σ = = 4.24

So, S.E. of Mean =  = 1.73



N= 4

Sample size, n = 2

Possible samples, 4C2= 6

|  |  |  |
| --- | --- | --- |
| SNo. | Sample | Sample proportion, p |
| 1 | 1,2 | ½ |
| 2 | 1,3 | 0 |
| 3 | 1,4 | ½ |
| 4 | 2,3 | ½ |
| 5 | 2,4 | 1 |
| 6 | 3,4 | ½ |

|  |  |  |
| --- | --- | --- |
| Sample proportion, (p) | Frequency (f) | RELATIVE freq. |
| ½ | 4 | 4/6 |
| 1 | 1 | 1/6 |
| 0 | 1 | 1/6 |

Unbiased = true value

Population proportion of even number, P = 2/4 = ½

**E(p) =**

= ½ \*6/4 + 1\*1/6 + 0\*1/6 = ½ = P

An electronics company manufactures resistors that have a mean resistance of 100 ohms and a standard deviation of 10 ohms. The distribution of resistance is normal. Find the probability that a random sample of 25 resistors will have an average resistance less than 95 ohms.

Mean , μ = 100

Standard deviation, σ = 10

Sample size, n = 25

Sample average,

P( ӯ < 95) = ?

By using Central Limit theorem,

P(ӯ < 95) = P(Z < ) = P(Z< -2.5) = 0.0062

Degrees of freedom

Suppose we have n = 5 observations

Sum = 20

X1 = 5, X2= 3, X3= 2, X4= 3, X5= **7**

**d.f. is the number of observations we are free to choose.**

**If n is number of observations, degrees of freedom is n-1**

Standard deviation, t1=

Standard deviation, t2 =

Sample size n=3 {1,5,6} mean= 4

Sample size, n= 7{4, 5, 8 , 6, 3, 10, 9} mean= 6.42

Median, {3, 5, 8, 17, 20}

Mode

Measures of dispersion,

Range, QD, Standard deviation

An auditor for a large credit card company knows that, on average, the monthly balance of any given customer is 112 and the standard deviation is 56. If the auditor audits 50 randomly selected accounts, what is the probability that the sample average monthly balance is (a) below 100 (b) between 100 and 130?

Average monthly balance, μ = 112

S.D. σ = 56

Sample size, n= 50

Sample average = ӯ

By Using CLT,

P(ӯ < 100) = P(Z< ) = P(Z < -1.52) =

P(100< ӯ < 130) =

An electronics company manufactures resistors that have a mean resistance of 100 ohms and a standard deviation of 10 ohms. The distribution of resistance is normal. Find the probability that a random sample of 25 resistors will have an average resistance less than 95 ohms.

mean resistance, μ = 100 ohms

standard deviation, σ = 10 ohms

sample size, n = 25

P(ӯ< 95)=?

Using Central Limit Theorem,

P(ӯ< 95)=P(Z<) = P(Z<-2.5) = use Z score table

A quality control manager needs to estimate the average hours of life of light bulbs. The population standard deviation is known to be 100 hours. A random sample of 64 light bulbs indicated a sample average life of 350 hours. Set up 95% and 99% confidence interval of true average life of bulbs.

Given,

Population standard deviation, σ = 100 hrs

Sample size, n = 64

Sample mean, ӯ = 350 hrs

For 95% confidence level, Zα/2 = 1.96

Now, LCL = Ӯ - Zα/2 \*σ/√n =

UCL = Ӯ + Zα/2 \*σ/√n =

The 95% confidence interval is

(LCL, UCL)

For 99% confidence level, Zα/2 = 2.575

Now, LCL = Ӯ - Zα/2 \*σ/√n =

UCL = Ӯ + Zα/2 \*σ/√n =

The 99% confidence interval is (LCL, UCL)

The weights of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2 and 9.6 ounces. Find a 95% confidence interval for the mean of all such containers, assuming an approximate normal distribution

Sample size, n= 7

Here, Ӯ = 10

Sample s.d., s = 0.2828

For 95% confidence level, tα/2,n-1 = 2.447

Lower Confidence Limit (LCL) = Ӯ - tα/2,n-1 \*s/√n =

Upper Confidence Limit (UCL) = Ӯ + tα/2,n-1 \*s/√n =

(a……………b)>>>>>>>>>>>>> 90%

(c…………...………d)>>>>>>>>> 95%

(e………………………………..f)>>>>> 99%

A quality control manager needs to estimate the **average** hours of life of light bulbs. The population standard deviation is known to be 100 hours. A random sample of 64 light bulbs indicated a sample average life of 350 hours. Set up 95% and 99% confidence interval of true average life of bulbs.

Given,

Sample size, n= 64

Standard deviation of life, σ = 100

Sample mean, = 350 hours

For Confidence level 95%, Zα/2= 1.96

LOWER confidence limit (LCL) = - Zα/2 \*σ/√n =

Upper confidence limit, (UCL)= + Zα/2 \*σ/√n =

So, the confidence interval is (LCL, UCL) = (a, b)

For Confidence level 99%, Zα/2= 2.575

LOWER confidence limit (LCL) = - Zα/2 \*σ/√n =

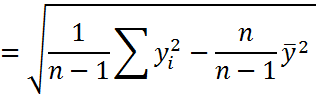
Upper confidence limit, (UCL)= + Zα/2 \*σ/√n =

So, the confidence interval is (LCL, UCL)=

The weights of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2 and 9.6 ounces. Find a 95% confidence interval for the mean of all such containers, assuming an approximate normal distribution.

Sample size, n= 7

Sample mean, = 10

Sample s.d. , s = = 0.2828

tα/2, n-1 = 2.447

LCL = Ӯ - tα/2,n-1 \*s/√n =

UCL = Ӯ +- tα/2,n-1 \*s/√n =

So, the confidence interval is (LCL, UCL)

11. For city 1,

Sample size, n1 = 90

Sample mean of income, = 28,520

Sample s.d., s1 = 1510

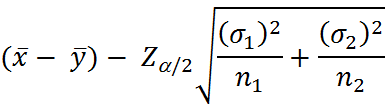
For city 2,

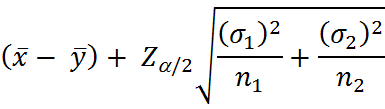
Sample size, n2 = 60

Sample mean of income, = 27,210

Sample s.d., s1 = 950

For 99% confidence level, Zα/2=2.575

Now, LCL =  = 792.58

UCL = = 1827.41

12. for 1st aluminum spar,

Sample size, m = 10

Sample tensile strength = 87.6

Sample standard deviation, s1= 1

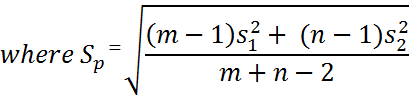
For 2nd aluminum spar,

Sample size, n = 12

Sample tensile strength =74.5

Sample standard deviation, s2= 1.5

For 90% confidence interval, t α/2, m+n-2 = 1.725

Sp= Pooled standard deviation 

Sp = Pooled (combined) standard deviation of two samples

12. For first type of aluminum spar,

Sample size, m = 10

Sample mean of tensile strength, = 87.6

Sample S.D. s1= 1

For second type of aluminum spar,

Sample size, n= 12

Sample mean of tensile strength , =74.5

Sample s.d S2 = 1.5

Now, for 90% confidence level, tα/2, m+n-2 = 1.725

5. Sample size, n= 500

Number of families who subscribe to HBO = 340

Sample proportion of families who subscribe to HBO, p =340/500 = 0.68

Now, q = 1-p = 0.32

For 95% confidence level, = 1.96

LCL = p - Zα/2 =0.639

UCL = p + Zα/2 = 0.721

In 40 tosses of a coin, 24 heads were obtained. Find 95% confidence limit for proportion of heads.

Given,

No. of toss, n = 40

Heads obtained, x = 24

Sample proportion of heads, p = 24/40= 3/5

Sample proportion of tails, q = 1- 3/5 = 2/5

For 95% confidences level, Zα/2 = 1.96

Lower confidence limit = p - Zα/2 =0.448

Upper confidence limit = p + Zα/2 = 0.752

The 95% Confidence interval for true proportion of heads = (0.448, 0.752)

A random sample of 500 adult residents of Town A found that 385 were in favor of increasing the highway speed limit to 75mph, while another sample of 400 adult residents of Town B found that 267 were in favor of the increased speed limit. Construct 95% confidence interval on the **difference in the two proportions**

For town A,

Sample size, m = 500

Number of respondents who are in favor of increasing highway speed limits = 385

Sample proportion in town A who are in favor of increasing highway speed limits, p1 = 385/500

For town B,

Sample size, n = 400

Number of respondents who are in favor of increasing highway speed limits = 267

Sample proportion in town A who are in favor of increasing highway speed limits, p2 = 267/400

Local people bought some CFL bulbs in order to save electricity. Upon testing those bulbs, they found that each CFL has a mean life of 10 months with a standard deviation of 1 month.

a. Find the probability that 36 CFL bulbs last at least 31 years. [Ans: 0.0228]

b. How many CFL bulbs should be bought so that the buyer can be 95% sure that CFL will last 10 years? [Ans: 12.7313]

Average life of CFL, μ= 10 months

Standard deviation, σ= 1 month

let xi be the life time of ith CFL bulb

then total life of 36 bulbs be

Sn = x1 + x2 + ………+ xn

By CLT,

Z = is standard normal variable

Prob. that 36 CFL bulbs last at least 31 years = 372 months

1. P(S36 ≥ 372) = ???

P(Z ≥ ) =P(Z≥ 2)= ……………..

1. Suppose ‘n’ bulbs is purchased so that the buyer is 95% confident that it will last 10 years = 120 months

Now, sum of life of n bulbs be Sn = x1+x2+……..+xn follows normal distribution according to Central Limit Theorem.

According to question,

P(Sn ≥ 120) = 0.95

P(Z ≥ ) = 0.95

P (Z < ) = 0.05

From Z score table,

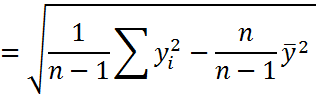
= - 1.645

Solve for n……

The weights of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2 and 9.6 ounces. Find a 95% confidence interval for the mean of all such containers, assuming an approximate normal distribution.

Sample size, n= 7

Sample mean, = = 10

Sample standard deviation, s =  = 0.2828

For 95% confidence level, tα/2, n-1 =2.447

|  |  |
| --- | --- |
| Large sample (n≥30 | Small sample, n<30 |
| Single mean (population s.d.σ is known) | Single mean |
| Single mean (population s.d.σ is unknown) | Difference mean |
| Difference mean |  |
| Proportion estimation, P |  |
| Difference of proportion, P1-P2 |  |

Suppose that we want to estimate the true proportion of defectives in a very large shipment of adobe bricks, and that we want to be at least 95% confident that the error is at most 0.04. How large a sample will be need if

(a) We have no idea what the true proportion might be. (601)

(b) We know that the true proportion does not exceed 0.12. (254)

For 95% confidence level, Zα/2 = 1.96

Maximum error of estimate, E = 0.04

1. P= 0.5, Q = 0.5
2. P= 0.12, Q = 0.88

μ = 50 (Assumption about population parameter) WRONG

= **48, 51, 52. …….**

= 80, 85, 30, 20….

= 40

Is the difference between μ and are significant or not?

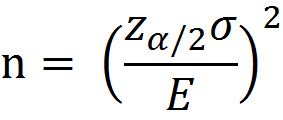
|  |  |  |  |
| --- | --- | --- | --- |
| **Large sample (n≥ 30)** | | **Small sample (n< 30)** | |
| **One sample** | **Two sample** | **One sample** | **Two sample** |
| Mean (σ known) | Difference of two means | Mean (t) | Difference of two sample means (t) |
| Mean (σ unknown) |  |  |  |
| Proportion | Difference of two proportions |  |  |
|  |  |  |  |

A researcher wants to determine the mean time required to complete certain job so that he may be 95% confident that mean may remain within  2 days of the true mean. As per available record, the population variance is 64 days. How large should the sample be for his study?

For 95% confidence, Zα/2 = 1.96

E, maximum error of estimate = 2

σ = 8

So, the sample size required = = 62

Suppose that we want to estimate the true proportion of defectives in a very large shipment of adobe bricks, and that we want to be at least 95% confident that the error is at most 0.04. How large a sample will be need if

(a) We have no idea what the true proportion might be. (601)

(b) We know that the true proportion does not exceed 0.12. (254)

For 95% confidence, Zα/2 = 1.96

E, maximum error of estimate = 0.04

1. P=Q= 0.5
2. P=0.12, q= 1-P = 0.88

μ = 50 (Assumption about population parameter)

Random sample= n

= **48, 45, 48, 51, 52 …….49, 45, 53**

= 80, 85,70…20.25.30

= 45

Is the difference between μ and are significant or not?

H0: μ =100

The 95 operators were trained to use the new machines averaged 7.2 hours before achieving a satisfactory performance. Their sample standard deviation was 4.02 hours. Previous experience shows that the operators used the old machines averaged 8.1 hours before achieving a satisfactory performance. At the 0.01 significance level, should we conclude that the new machines are **easier** to operate?

Given,

Sample average of new machines, = 7.2 hours,

Sample size, n = 95,

Sample standard deviation, s = 4.02

1. Null Hypothesis (H0): μ= 8.1

Alternative hypothesis (H1): μ < 8.1

1. Level of significance, α = 0.01
2. Test Statistic, Z = = -2.182
3. Critical value, Zα = - 2.33
4. Decision: Since - Z > - Zα, there is no significant difference between time operation. So, we accept H0 and reject H1. We conclude that new machine is not easier to operate.

29. The director at a university advises parents of new students about the cost of textbooks. A sample of 100 students enrolled in the university indicates a sample mean of cost Rs 3150 with sample standard deviation of Rs 43. Using 0.10 level of significance, is there evidence that the population average is **above** Rs 3000?

Sample size, n = 100

Sample mean, = Rs 3150

Sample standard deviation, s = Rs 43

1. Null Hypothesis (H0): μ= 3000

Alternative hypothesis (H1): μ > 3000

1. Level of significance, α = 0.1
2. Test Statistic, Z = = 34.88
3. Critical value, Zα = 1.28
4. Decision: Since Z > Zα, there is significant difference in the cost of text books, so, we reject H0 and accept H1. **(THERE IS NO ENOUGH EVIDENCE TO SUPPORT NULL HYPOTHESIS)** We can conclude that cost of text books are above Rs 3000.

30. In the past a machine has produced washers having a mean thickness of 1.250 mm. To determine whether the machine is in proper working condition, a sample of 100 washers is chosen for which the mean thickness is 1.325 mm and s.d. is 0.08 mm. Test the hypothesis that the machine is in proper working condition using a level of significance 0.05.

Given,

Sample size, n = 100

Sample mean, = 1.325 mm

Sample S.d., s = 0.08 mm

1. Null Hypothesis (H0): μ= 1.250 mm

Alternative hypothesis (H1): μ ≠ 1.250 mm

1. Level of significance, α = 0.05
2. Test Statistic, Z = = 9.375
3. Critical value, Zα/2= ± 1.96
4. Decision: Since Z > Zα/2, there is significant difference in the thickness of washers, so, we reject H0 and accept H1. We can conclude that machine is not in proper working condition.

33. In a certain factory, there are two independent process of manufacturing same item. The average weight in a sample of 250 items produced from one process is found to be 120 gram with standard deviation of 12 gram while the corresponding figure in a sample of 400 items from the other process are 124 and 14. Find the standard error of the difference of means and also **test whether two mean weight differ significantly or not** at 10% level of significance.

Given,

For 1st process,

Sample size, m = 250

Sample mean, = 120 gram

Sample standard deviation, sx = 12 gram

For 2nd process,

Sample size, n = 400

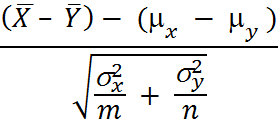
Sample mean, = 124 gram

Sample standard deviation, sy = 14 grams

Let μx and μy be the mean weight of two process.

1. Null Hypothesis, (H0): μx - μy = 0

Alternative hypothesis, H1: μx - μy ≠ 0

1. Level of significance, α = 10% = 0.1
2. Critical value, Zα/2 = ± 1.645
3. Test statistic, Z = = - 3.87
4. Decision: Since |Z| > | Zα/2,| the difference is significant, so, there is no sufficient evidence to accept the H0. The mean weight difference from two process is significant.

36. A manufacturer intends that his electric light bulbs have a life of 1000 hours. He tests a sample of 20 bulbs, drawn at random from a batch and discovers that the mean life of the sample bulbs is 990 hours with a standard deviation of 22 hours. Does this signify that the batch is not up to the standard? Use  = 0.05

Given,

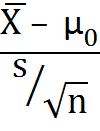
Sample size, n= 20

Sample mean, = 990 hours

Sample s.d., s = 22 hours

1. Null hypothesis, H0: μ = 1000 hrs

Alternative hypothesis, H1: μ < 1000 hrs

1. Level of significance, α= 0.05
2. Critical value, tα, n-1 = 1.729
3. Test statistic, t = = - 2.03
4. Decision: Since |t| > | tα, n-1|, the difference is significant, so we reject the null hypothesis and accept alternative hypothesis. So, we conclude that batch is not up to the standard.

40. The height of six randomly chosen sailors are 68, 65, 68, 69, 71 and 72 inches. Those of 10 randomly chosen soldiers are 61, 65, 66, 69, 70, 70, 71, 62, 72 and 70 inches. Can we assume that data support the fact that sailors are taller on the average than soldiers do?

Let μx and μy be the mean height of sailors and soldiers respectively.

For sailors,

Sample size, m = 6  
sample mean of height of sailors, = 68.83

Sample standard deviation, sx = 2.591

For soldiers,

Sample size, n = 10

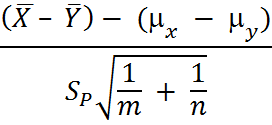
Sample mean of height of soldiers, = 67.6

Sample standard deviation, sy = 3.864

Pooled standard deviation, Sp = 3.463

1. Null hypothesis, H0: μx - μy = 0

Alternative hypothesis, H1: μx > μy

1. Level of significance, α = 0.01
2. Test statistic, t = = 0.687
3. Critical value, tα, m+n-2 = 2.624
4. Decision: Since t < tα, m+n-2, we do not have sufficient evidence to reject H0. The difference between height of sailors and soldiers are not significant.

42. The following random samples are measurement of the heat producing capacity (in millions of calories per ton) of specimens of coal from two mines:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mine 1 | 8260 | 8130 | 8350 | 8070 | 8340 | - |
| Mine 2 | 7950 | 7890 | 7900 | 8140 | 7920 | 7840 |

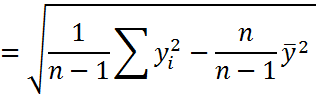
Use the 0.01 level of significance to test whether the difference between means of these two samples is significant

Given,

For mine 1,

Sample size, m= 5

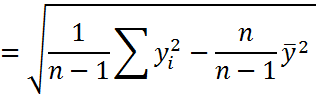
Sample average heat producing capacity, = 8230

Sample s.d. sx = = 125.499

For mine 2,

Sample size, n = 6

Sample average heat producing capacity, = 7940

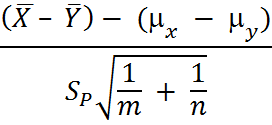
Sample s.d. sy = = 104.49

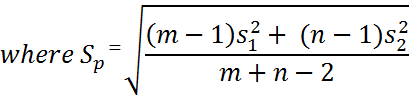
Let μx be average heat producing capacity from Mine 1

μy be average heat producing capacity from Mine 2

1. Null hypothesis, H0: μx - μy = 0

Alternative hypothesis, H1: μx - μy ≠ 0

1. Level of significance, α = 0.01
2. Test statistic, t = = 4.19

= 114.3

1. Critical value, tα/2, m+n-2 = 3.250
2. Decision: since t > tα/2, m+n-2, we do not have sufficient statistical evidence to accept H0, so we accept H1. The difference of heat producing capacity of two mines is significantly different.

44. An experiment is run to determine the effect of a new type drug on blood pressure. Ten persons have their blood pressure measured before and after the drug is given. The result of the experiment are as follows:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Blood Pressure | | | | | | | | | | |
| Persons | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Before | 116 | 118 | 120 | 124 | 128 | 130 | 131 | 134 | 136 | 137 |
| After | 119 | 124 | 126 | 128 | 121 | 135 | 137 | 138 | 139 | 135 |

Test the hypothesis the new drug does not raise blood pressure at 0.05 level of significance

Given,

Number of pairs, n = 10

Let μx and μy be the average blood pressure before and after using drug.

Now, sample mean of blood pressure before using drug, = 127.4

Sample mean of blood pressure after using drug, = 130.2

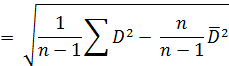
**Differences of each pair,**

D1 = 116-119= - 3

………………………….,

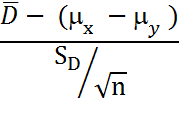
D10= 137-135 = 2

= - 2.8

SD = 4.184

1. Null hypothesis, H0: μx - μy = 0

Alternative hypothesis, H1: μx - μy > 0

1. Level of significance, α = 0.05
2. Test statistic, t = = -2.116
3. Critical value, tα, n-1 = 1.833
4. Decision: since |t| > |tα, n-1|, we do not have evidence to accept H0, so we accept H1. i.e. the difference is significant and new drug do not raise blood pressure.

Six samples of each of four types of cereal grain grown in a certain region were analyzed to determine thiamin content, resulting in the following data (mg/g):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Wheat | 5.2 | 4.5 | 6.0 | 6.1 | 6.7 | 5.8 |
| Barley | 6.5 | 8.0 | 6.1 | 7.5 | 5.9 | 5.6 |
| Maize | 5.8 | 4.7 | 6.4 | 4.9 | 6.0 | 5.2 |
| Oats | 8.3 | 6.1 | 7.8 | 7.0 | 5.5 | 7.2 |

Does this data suggest that at least two of the grains differ with respect to true average thiamin content? Use a 0.05 level of significance

Let μ1, μ2, μ3 and μ4 be the average thiamin content in Wheat, Barley, Maize and Oats respectively.

Number of groups, k = 4 (i = 1, 2, 3, 4)

= mean of ith group

= 5.72

= 6.6

= 5.5

= 6.98

T = Total of all observations = 148.8

N = total number of observations = 24

= Grand mean = 6.2

Null Hypothesis, H0: μ1 =μ2 =μ3 = μ4

Alternative Hypothesis, H1: μ1 ≠μ2 ≠μ3 ≠ μ4

Level of significance, α= 0.05

Test Statistic: F = 

Steps of calculation:

1. Sum of squares between treatment (SSTr)= n1(ӯ1- ӯ**..**)2 + n2(ӯ2- ӯ**..**)2 …..+ nk(ӯk- ӯ**..**)2 = 8.93
2. Total sum of squares (TSS) = **ΣΣ**yij 2 – T2/N = 18.87
3. Sum of squares within treatment (SSW)= 9.94

MSTr = SSTr/k-1 = 2.98

MSW = SSW/ N-k = 0.497

F = 5.99

Critical Value, Fα, k-1, N-k =3.0984

Decision: Since, F > Fα, k-1, N-k, there is significant difference in average thiamin content of the cereal grains. So we do not have evidence to accept H0.

∑∑yij2 = 946.68

Six samples of each of four types of cereal grain grown in a certain region were analyzed to determine thiamin content, resulting in the following data (mg/g):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Wheat | 5.2 | 4.5 | 6.0 | 6.1 | 6.7 | 5.8 |
| Barley | 6.5 | 8.0 | 6.1 | 7.5 | 5.9 | 5.6 |
| Maize | 5.8 | 4.7 | 6.4 | 4.9 | 6.0 | 5.2 |
| Oats | 8.3 | 6.1 | 7.8 | 7.0 | 5.5 | 7.2 |

Does this data suggest that at least two of the grains differ with respect to true average thiamin content? Use a 0.05 level of significance.

Let μ1, μ2, μ3, μ4 be the average thiamin content of Wheat, Barley, Maize and Oats respectively.

Number of groups (treatments), k = 4 (i = 1,2,3,4)

Mean of each group,

1 = sample mean of thiamin content in Wheat = 5.72

2 = sample mean of thiamin content in barley = 6.6

3 = sample mean of thiamin content in maize = 5.5

4 = sample mean of thiamin content in oats = 6.98

Number of samples in each group, n1=n2=n3=n4= 6

T = 148.74

N= 24

= grand mean= T/N = 6.2

1. Null hypothesis, H0: μ1 = μ2 = μ3 = μ4

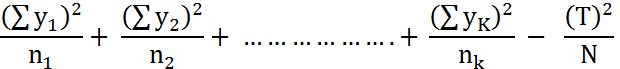
Alternative hypothesis, H1: μ1 ≠ μ2 ≠ μ3 ≠ μ4

1. Level of significance, α = 0.05
2. Test statistic, F = 

Steps of Calculations

Sum of squares between group (treatment) (SSTr) = n1(ӯ1- ӯ**..**)2 + n2(ӯ2- ӯ**..**)2 …..+ nk(ӯk- ӯ**..**)2

= 8.93

= 8.98

* Total sum of squares, TSS = **ΣΣ**yij 2 – T2/N = 24.12
* Sum of squares within group ,SSW = TSS – SSTr =15.19
* Mean sum of squares between group, MSTr = = 2.976
* Mean sum of squares within group, MSWr = =0.7595
* F = = 3.918

1. Critical value,Fα, k-1,n-k = 3.0984
2. Decision: since F > Fα, k-1,n-k, we do not have sufficient statistical evidence to accept H0, there is significant difference between average thiamin content in four given cereal grain.

It is thought that the proportion of defective items produced by a particular machine is 0.1. A random sample of 100 items is inspected and found to contain 16 defective items. Does this provide evidence, at the 5% level, that the machine is producing more defective items than expected?

Given,

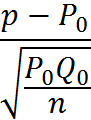
Sample size, n= 100

Defective numbers, x = 16

Sample proportion of defectives, p = 16/100 = 0.16

1. Null hypothesis, H0: P = 0.1

Alternative hypothesis, H1: P > 0.1

1. Level of significance, α= 0.05
2. Critical value, Zα = 1.645
3. Test statistic, Z = = 2
4. Decision: Since |Z| > | Zα, n-1|, the difference is significant, so we do not have sufficient evidence to accept the null hypothesis. So, we conclude that machine is producing more than 10% defective items.

A coin is tossed 200 times and 115 heads obtained. Is there evidence, at the 1% level, that the coin is biased towards heads?

Sample proportion of head, p = 115/200 =

H0: P= 0.5

H1:P > 0.5

In a sample of 600 men from a certain city, 450 are found to be smokers. In a sample of 900 from another city 450 are found to be smokers. Do the data indicate that the two cities are significantly different with respect to prevalence of smoking habit among men?

For 1st city,

Sample of men, n1 = 600

Number of smokers, x = 450

Sample proportion of smokers in this city, p1 = 450/600 =

For 2nd city,  
sample size, n2 = 900

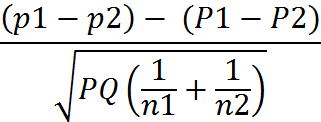
Number of smokers, y = 450

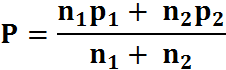
Sample proportion of smokers, p2 = 450/900

Let P1 and P2 be the true proportion of smokers in 1st and 2nd city.

1. Null Hypothesis, H0 : P1-P2= 0

Alternative hypothesis, H1: P1 ≠ P2

1. Level of significance, α = 0.05
2. Test statistic, Z= = 9.682

Where, , Q = 1-P

1. Critical Value, Z α/2 = ± 1.96
2. Decision: Since, Z > Z α/2, there is significant difference in proportion of smokers in the two cities. We do not have sufficient evidence to accept the null hypothesis.

**A dice is thrown 60 times with following results. Test at 5% level of significance if the dice is unbiased.**

Null Hypothesis, H0: There is no significant difference between observed and expected frequencies.

(DICE IS UNBIASED i.e.prob. of getting each face is 1/6)

H1: There is significant difference between observed and expected frequencies.

(DICE IS BIASED. i.e. prob. of getting each face is not 1/6)

Here no. of categories, k = 6

|  |  |  |  |
| --- | --- | --- | --- |
| face | Observed Frequency (O) | Expected Frequencies (E) |  |
| 1 | 8 | 60\*1/6 = 10 | 0.4 |
| 2 | 7 | 10 | 0.9 |
| 3 | 12 | 10 | 0.4 |
| 4 | 8 | 10 | 0.4 |
| 5 | 14 | 10 | 1.6 |
| 6 | 11 | 10 | 0.1 |
| Total | 60 | 60 | 3.8 |

Level of significance, α= 0.05

Test Statistic, χ2 = 3.8

Critical value, χ20.05, k-1 = 11.071

Decision: since χ2 <χ20.05, k-1 we do not have sufficient evidence to reject H0. So the dice is considered to be unbiased.

17. An observed frequency distribution is as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. of success (x) | Frequency (O) | Binomial prob.  P(X=x)= C(n,x)pxqn-x | Expected frequency (E) |  |
| 0 | 89 | 8/27 | 300\*8/27= 88.89 | 0.000136 |
| 1 | 133 | 4/9 | 300\*4/9= 133.33 | 0.000816 |
| 2 | 52 | 2/9 | 300\*2/9= 66.67 | 3.227 |
| 3 | 26 | 1/27 | 300\*1/27= 11.11 | 19.95 |
| Total | 300 | 1 |  |  |

Given,

For binomial distribution,

Number of trials n = 3

Probability of success, p = 1/3, q= 2/3

Number of categories, k = 4

Null hypothesis, H0: The given distribution fits into binomial distribution with n=3 and p=1/3

Alternative hypothesis, H1: The given distribution does not fit into binomial distribution with n=3 and p=1/3

Level of significance, α = 0.05

Test Statistic, χ2= 23.177952

Critical value, χ2α, k-1= 7.815

Decision: Sinceχ2 > χ2α, k-1, the given distribution does not fit into binomial distribution.

12. Suppose that a random sample of men and women indicated their view on a certain proposal as shown in table (2 by 3)

Test the statement that there is no difference in opinion between men and women, i.e. the response is independent of the sex of the person interviewed at a significance level of  = 0.05

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observed Value** | | | | |
|  | in favor | opposed | undecided | |
| women | 118 | 62 | 25 | **205** | |
| men | 84 | 78 | 37 | **199** | |
| Total | **202** | **140** | **62** | **404** | |

Null Hypothesis, H0: The gender and opinion are independent.

Alternative Hypothesis H1: The gender and opinion are not independent.

Level of significance, α= 0.05

Calculations:

Expected Values

|  |  |  |  |
| --- | --- | --- | --- |
|  | In favor | Opposed | Undecided |
| Women | E(118) = (205\*202)/404 = 102.5 | E(62) = (205\*140)/404  = | E(25) = (205\*62)/404 |
| Men | E(84) = (199\*202)/404 | E(78) = 199\*140/404 | E(37) = (199\*62)/404 |

Test Statistic, χ2= + ……………………………………………. = 9.79

Critical value, χ2α, (r-1)\*(c-1)= 5.991

r=number of row= 2

c= number of columns = 3

decision: since χ2> χ2, (r-1) χ2\*(c-1) we conclude that gender and opinion are not independent.