**Q.** Following are the marks obtained by 13 students in business

statistics. Obtain mean, median, mode,25th percentile, first

quartile and third quartile using excel.

|  |  |
| --- | --- |
| formula |  |
| n | 13 |
| Σx | 734 |
| mean | 56.46154 |
| median | 54 |
| mode | 52 |
| 25th per | 48 |
| Q1 | 48 |
| Q3 | 60 |

|  |
| --- |
| Marks |
| 24 |
| 27 |
| 36 |
| 48 |
| 52 |
| 52 |
| 54 |
| 55 |
| 56 |
| 60 |
| 85 |
| 90 |
| 95 |

**Remarks:**

The metrics indicate that marks obtained by students is not symmetric,

rather it is positively skewed. The first quartile and 25th percentile are

both 48, which means 25% of the students scored below 48. While the

third quartile is 60, which shows that 75% of students scored below 60

marks.

**Q.** Box and Whisker plot

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | difference |
| 0 | xmin | 9 | 9 |
| 1 | Q1 | 16 | 7 |
| 2 | md | 25 | 9 |
| 3 | Q2 | 30 | 5 |
| 4 | xmax | 50 | 20 |

|  |
| --- |
| X |
| 9 |
| 33 |
| 47 |
| 15 |
| 22 |
| 17 |
| 18 |
| 45 |
| 50 |
| 26 |
| 25 |
| 27 |
| 9 |
| 27 |
| 10 |

**Remarks:**

The box and whisker plot clearly demonstrates that the

Distribution is positively skewed as the whisker of positive side is

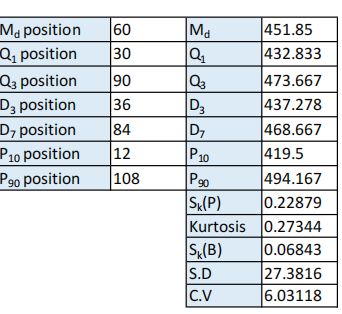
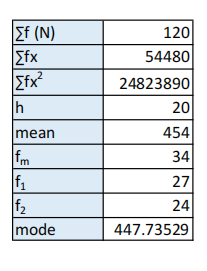
longer than that of the negative side.

Q. Compute median, Q1, Q3, D3, D7, P90, P10, Karl Pearson's

skewness (SkP), Kurtosis, Bowley's skewness (SkB), S.D, and C.V

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| milage | 400-419 | 420-439 | 440-459 | 460-479 | 480-499 | 500-519 |
| frequency | 12 | 27 | 34 | 24 | 15 | 8 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| LCB | UCB | mid(x) | frequency | fx | CF | X2 | FX2 |
| 399.5 | 419.5 | 409.5 | 12 | 4914 | 12 | 167690.3 | 2012283 |
| 419.5 | 439.5 | 429.5 | 27 | 11596.5 | 39 | 184470.3 | 4980697 |
| 439.5 | 459.5 | 449.5 | 34 | 15283 | 73 | 202050.3 | 6869709 |
| 459.5 | 479.5 | 469.5 | 24 | 11268 | 97 | 220430.3 | 5290326 |
| 479.5 | 499.5 | 489.5 | 15 | 7342.5 | 112 | 239610.3 | 3594154 |
| 499.5 | 519.5 | 509.5 | 8 | 4076 | 120 | 259590.3 | 2076722 |
|  |  |  | 120 | 54480 | 120 | 1273842 | 24823890 |



**Remarks:** The metrics show that the distribution is positively skewed and is

leptokurtic. The C.V of the distribution is only 6.03%, which indicates that

the distribution is relatively uniform.

Q. Draw a scatter diagram of daily rainfall and particular level in Kathmandu valley using excel

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Daily rainfall(x) | particular level(y) | XY | X2 | Y2 |
| 4.1 | 122 | 500.2 | 16.81 | 14884 |
| 4.3 | 117 | 503.1 | 18.49 | 13689 |
| 5.7 | 112 | 638.4 | 32.49 | 12544 |
| 5.4 | 114 | 615.6 | 29.16 | 12996 |
| 5.9 | 110 | 649 | 34.81 | 12100 |
| 5 | 114 | 570 | 25 | 12996 |
| 3.6 | 128 | 460.8 | 12.96 | 16384 |
| 1.9 | 137 | 260.3 | 3.61 | 18769 |
| 7.3 | 104 | 759.2 | 53.29 | 10816 |
| 43.2 | 1058 | 4956.6 | 226.62 | 125178 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| n | 9 |  | y | 85.07644 |
| NXY | 44609.4 |  | ∑x2.n-∑x2 | 13.16586 |
| sumX\*sumY | 45705.6 |  | r | -0.97866 |
| n\*sumX2 | 2039.58 |  | se | 0.014076 |
| (sumX)2 | 1866.24 |  | b | -6.32399 |
| n\*sumy2 | 1126602 |  | a | 147.9107 |
| (sumy) | 1119364 |  | y=147.91-6.32X | |

|  |  |  |
| --- | --- | --- |
|  | Column 1 | Column 2 |
| Column 1 | 1 |  |
| Column 2 | 0.988602 | 1 |

|  |  |
| --- | --- |
| *Regression Statistics* | |
| Multiple R | 0.978658 |
| R Square | 0.957772 |
| Adjusted R Square | 0.95174 |
| Standard Error | 2.202613 |
| Observations | 9 |

**Remarks:**

The distribution has a highly negative correlation coefficient. This

means that the dependent variable (y) will decrease drastically even

the when independent variable (x) increases slightly.

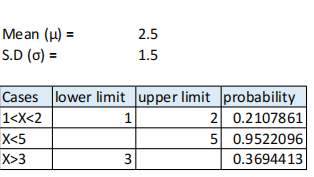
This conclusion can also be drawn with the help of the above scatter

plot. As we move towards higher values of rainfall, the particulate

level decreases.

**Q.** A random variable follows normal distribution with parameter µ=2.5

and σ =1.5. Find the probabilities of 1<X<2 , X<5, and X>3.



**Remarks:**

The parameters for normal distribution are x (value), µ (mean), and

σ (S.D).

By using normal distribution, probability for required cases are

calculated, with the probability of case (x<5) being the highest, i.e.,

0.9522.

**Q.** Pancha kanya group of industries produces tube. The length of a

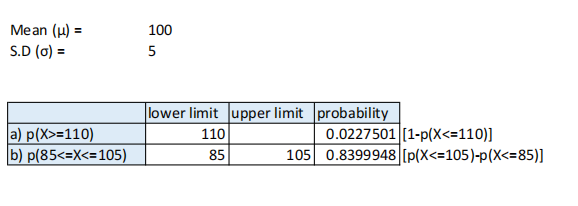
tube is found to be normally distributed with a population mean 100

and standard deviation 5.

(a) Calculate probability that a random sample of one tube will have a

length of at least 110.

(b) p(85<=X<=105).



Q. Find the expected mean

|  |  |  |  |
| --- | --- | --- | --- |
| x | p(X=x) | x.p(X=x) | x^2\*p(X=x) |
| 0 | 0.1 | 0 | 0 |
| 1 | 0.2 | 0.2 | 0.2 |
| 2 | 0.45 | 0.9 | 1.8 |
| 3 | 0.15 | 0.45 | 1.35 |
| 4 | 0.05 | 0.2 | 0.8 |
| 5 | 0.05 | 0.25 | 1.25 |
| 15 | 1 | 2 | 5.4 |

|  |  |
| --- | --- |
| mean | 15 |
| E(x)^2 | 5.4 |
| var | 1.4 |
| E(4x+7) | 15 |
| v(4x+7) | 22.4 |

**Remark:**

**Q.** Calculate spearman's correlation coefficient between statistics

and mathematics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A | B | Rank A | Rank B | d | d2 |
| 67 | 31 | 9 | 5 | 4 | 16 |
| 41 | 29 | 2 | 4 | -2 | 4 |
| 52 | 70 | 4 | 9 | -5 | 25 |
| 60 | 79 | 7 | 10 | -3 | 9 |
| 42 | 21 | 3 | 2 | 1 | 1 |
| 39 | 19 | 1 | 1 | 0 | 0 |
| 56 | 68 | 6 | 8 | -2 | 4 |
| 61 | 25 | 8 | 3 | 5 | 25 |
| 53 | 40 | 5 | 7 | -2 | 4 |
| 69 | 37 | 10 | 6 | 4 | 16 |
| n=10 |  |  |  |  | 104 |

|  |  |
| --- | --- |
| n(n^2-) | 990 |
| r(s)=1-6∑d2/n(n^2-1) | |
| 0.369697 |  |

**Remarks:**

Spearman's rank correlation coefficient is 0.369697. It indicates that

there is a low degree positive correlation between ranks of statistics

and math.

**Q.** Calculate spearman's correlation coefficient between statistics

and mathematics

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Anger | 6 | 7 | 5 | 21 | 13 | 5 | 13 | 14 |
| Vigor | 30 | 23 | 29 | 22 | 19 | 19 | 28 | 19 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Anger | Vigor | r1 | r2 | d | d^2 |
| 6 | 30 | 6 | 1 | 5 | 25 |
| 7 | 23 | 5 | 4 | 1 | 1 |
| 5 | 29 | 7.5 | 2 | 5.5 | 30.25 |
| 21 | 22 | 1 | 5 | -4 | 16 |
| 13 | 19 | 3.5 | 7 | -3.5 | 12.25 |
| 5 | 19 | 7.5 | 7 | 0.5 | 0.25 |
| 13 | 28 | 3.5 | 3 | 0.5 | 0.25 |
| 14 | 19 | 2 | 7 | -5 | 25 |
|  |  |  |  |  | 110 |

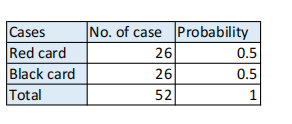
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 13 repeated 2 times | | m1 | 2 | 0.5 |
| 5 repeated 2 times | | m2 | 2 | 0.5 |
| 19 repeated 3 times | | m3 | 3 | 2 |
| n(n^2 -1) | 504 |  |  |  |
| r(s) | -0.34524 |  |  |  |

**Remarks:**

Hence rank correlation is negative so it is low degree correlation between Angor and vigor.

Q. A card is drawn at random from a well shuffled pack cards.

What is the probability of getting (a) red card (b) black card?



**Remarks:**

The probability of getting red and blacks’ cards are equal to 0.5.

Hence, both types of cards are equally likely to be drawn from the

card shuffle

Q. You are given below the income distribution of 1000 person. Find the

probability that a person selected has (a) income below 2000 and (b) income

2500 or more

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| income | 0-500 | 500-1000 | 1000-1500 | 1500-2000 | 2000-2500 | 2500-3000 | 3000-3500 | N |
| no.person | 150 | 250 | 300 | 100 | 80 | 70 | 50 | 1000 |

|  |  |  |
| --- | --- | --- |
| income | no.person | probability |
| 0-500 | 150 | 0.15 |
| 500-1000 | 250 | 0.25 |
| 1000-1500 | 300 | 0.3 |
| 1500-2000 | 100 | 0.1 |
| 2000-2500 | 80 | 0.08 |
| 2500-3000 | 70 | 0.07 |
| 3000-3500 | 50 | 0.05 |

|  |  |  |
| --- | --- | --- |
| prob income below 2000 | | |
| 0.8 |  |  |
| prob income 2500 or more | | |
| 0.2 |  |  |

**Remarks:**

When a person is selected, the probability that the person has income below

2000 is 0.8 (very high) and income 2500 or more is 0.2. It clearly shows that

there are more people having low income.

Q. A problem in statistics is given to three students A,B,C whose

chance of solving it are 1/3, 1/4 and 1/5 respectively. Find the

probability that

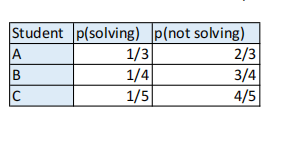
a. The problems will be solved

b. Only one of them can solve the problems

c. No. of them can solve the problems

d. A solve it but B and C cannot

e. All three students can solve the problems.



**Answers:**

a) => 3/5

60.00%

b) => 3/7

43.33%

c) => 2/5

40.00%

d) => 1/5

20.00%

e) => 0.016667

1.67%

.

**Remarks:**

The results show that the chances that the problem will be solved

when all three students try is **60%**, which is relatively high, while

the chances that all three students solve the problem is quite less,

i.e., **1.67%**.

The following table shows the survey regarding the employment status and gender in a sample of 209 management graduate.

|  |  |  |  |
| --- | --- | --- | --- |
|  | employed status | |  |
| gender | employed status | not | total |
| male | 83 | 28 | 111 |
| female | 64 | 34 | 98 |
|  | 147 | 62 | 209 |

what is the prob that a graduate chosen at random

a.is currently employed?

b.is a female and currently employed?

c.is female or currently employed?

d. suppose the graduate chosen is a female , what is the prob that she is currently employed ?

|  |  |
| --- | --- |
| a | 0.703349 |
| b | 1/3 |
| c | 0.346939 |
| d | 0.435374 |

**Remark:**

The probability of currently employed is 0.703349.

The probability of a female and is currently employed is 1/3.

The probability of female or currently employed is 0.346939.

The probability that she is currently employed is 0.435374

Q. A person shoots three rounds at target. The probability of hitting

a target successfully is 0.3. Compute the probability of hitting

targets where x= 0,1,2,3.

|  |  |
| --- | --- |
| n | 3 |
| p | 0.3 |
| q | 0.7 |

|  |  |
| --- | --- |
| probability | p(x) |
| 0 | 0.343 |
| 1 | 0.441 |
| 2 | 0.189 |
| 3 | 0.027 |

**Remarks:**

This problem can be solved by using binomial distribution as there

is probability of success given, which is constant for finite number

of trials he made.

The results clearly show that the probability for the person to hit

the target once is the highest, i.e., 0.441, which is still below 0.5.

Hence, he is relatively less likely to hit the target at least once.

On the other hand, the probability that the person will hit the

target in all his attempts is the minimum, i.e., 0.027.

Q. A random variable follows binomial distribution with parameter

n=6 and p=0.4. Find the probability of x=2, x<5, and x>3.

|  |  |
| --- | --- |
| n | 6 |
| p | 0.4 |
| q | 0.6 |

|  |  |  |
| --- | --- | --- |
| cases | X=x | probability |
| x=2 | 2 | 0.31104 |
| x<5 | 4 | 0.95904 |
| x>3 | 3 | 0.1792 |

**Remarks:**

The parameters for binomial distribution are x (no. of favorable

cases), n (total number of trials), and p (probability of success).

By using binomial distribution, probability for required cases are

calculated, with the probability of case (x<5) being the highest, i.e.,

0.95904.

Q. In a chips factory machine A, B and C manufactured respectively 25%,35% and 40% of the total of their output 5,4 and 2 percent defective bolts. A chips is drawn at random from the product and is found to be defective .What is the probability that it was manufactured by machine A,B and C?

|  |  |
| --- | --- |
| P(A) | 0.25 |
| p(B) | 0.35 |
| P(C) | 0.4 |
| p(D/A) | 0.05 |
| P(D/B) | 0.04 |
| P(D/C) | 0.02 |
|  |  |
|  |  |
| P(A/D) | 0.362319 |
| P(B/D) | 0.405797 |
| P(C/D) | 0.231884 |

Remark:

The probability that it was manufacture by machine A, B, C are 0.362319, 0.405797 and 0.231884 respectively.