Example 2.1: Calculate the harmonic mean, standard deviation and skewness coefficient for the following data. 25, 34, 43, 19, 36, 26, 38, 17, 25

Solution: $\sum x = 263, \ \ \bar{x} = 263/9 = 29.222, \ H.M = 9/(1/25 + ... + 1/25) = 26.73$ $\sum$

x	25	34	43	19	36	26	38	17	25	263
$x-\bar{x}$	-4.222	4.778	13.778	-10.222	6.778	-3.222	8.778	-12.222	-4.222	0.00
$(x-\bar{x})^2$	17.83	22.83	189.83	104.49	45.94	10.38	77.05	149.38	17.83	635.56
$(x-\bar{x})^3$	-75.73	109.06	2615.40	-1068.2	311.36	-33.46	676.32	-1825.8	-75.27	634.19

$$S_x = (635.56/8)^{0.5} = 8.913, \ m_3 = 634.19/9 = 70.47, \ C_s = \frac{9*9}{8*7} * \frac{70.47}{8.913^3} = 0.144$$

Example 2.2: 32 yearly mean discharge values of a stream (m^3/s) are given as follows. By taking class interval as 3 m^3/s, classify the data, obtain the frequency values and considering mid values of all classes, calculate the mean and standard deviation:

28 19 16 11 19 20 17 15 13 16 24 13 18 20 23 20

15 13 10 17 21 19 18 24 12 21 25 26 13 18 27 14

Solution: Mid values of each group are x_i and frequencies are calculated as $f_i = 100*x_i/32$

GROUP	10-12	13-15	16-18	19-21	22-24	25-27	28-30	$\sum$
N_i	3	7	7	8	3	3	1	32
f_i	9.375	21.875	21.875	25.000	9.375	9.375	3.125	100
x_i	11	14	17	20	23	26	29	-
$x_i N_i$	33	98	119	160	69	78	29	586
$(x_i-\bar{x})^2 * N_i$	160.4	130.2	12.1	22.8	65.9	177.3	114.2	683.875

$\bar{x}$ = 586/32 = 18.3125, S_x = $(683.875/32)^{0.5}$ = 4.620

Example 2.3: The following data are the mean annual precipitation heights (cm) of a city. Calculate the mean, standard deviation and skewness coefficient by

a. Not classifying the data.

68 45 78 74 54 98 87 74 90 75 79 67 80 83 85 92 82 73 87 82 73 90 85 77 73 78 85 95 74 79 88 87

Solution:

N_i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$		N_i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$
1	68	-11,28125	127,2666	-1435,73		17	82	2,71875	7,391602	20,09592
2	45	-34,28125	1175,204	-40287,5		18	73	-6,28125	39,4541	-247,821
3	78	-1,28125	1,641602	-2,1033		19	87	7,71875	59,5791	459,8762
4	74	-5,28125	27,8916	-147,303		20	82	2,71875	7,391602	20,09592
5	54	-25,28125	639,1416	-16158,3		21	73	-6,28125	39,4541	-247,821
6	98	18,71875	350,3916	6558,893		22	90	10,71875	114,8916	1231,494
7	87	7,71875	59,5791	459,8762		23	85	5,71875	32,7041	187,0266
8	74	-5,28125	27,8916	-147,303		24	77	-2,28125	5,204102	-11,8719
9	90	10,71875	114,8916	1231,494		25	73	-6,28125	39,4541	-247,821
10	75	-4,28125	18,3291	-78,4715		26	78	-1,28125	1,641602	-2,1033
11	79	-0,28125	0,079102	-0,02225		27	85	5,71875	32,7041	187,0266
12	67	-12,28125	150,8291	-1852,37		28	95	15,71875	247,0791	3883,775
13	80	0,71875	0,516602	0,371307		29	74	-5,28125	27,8916	-147,303
14	83	3,71875	13,8291	51,42697		30	79	-0,28125	0,079102	-0,02225
15	85	5,71875	32,7041	187,0266		31	88	8,71875	76,0166	662,7697
16	92	12,71875	161,7666	2057,469		32	87	7,71875	59,5791	459,8762
TOTAL							2 537	0.0000	3 692.469	-43 345.2

$$\bar{x} = \frac{\sum x}{N} = \frac{2537}{32} = 79.281, V_x = \frac{\sum (x - \bar{x})^2}{N} = \frac{3692.469}{32} = 115.39 = m_2, \quad S_x = \sqrt{V_x} = \sqrt{115.39} = 10.742,$$

$$m_3 = \frac{\sum (x - \bar{x})^3}{N} = \frac{-43355.2}{32} = -1354.95$$

$$C_s = \frac{N^2}{(N-1)(N-2)} * \frac{m_3}{m_2^{1.5}} = \frac{32^2}{31*30} * \frac{-1354.95}{115.39^{1.5}} = -1.204 < 0 \Rightarrow \text{Left skewed distribution}$$

b. Classifying the data into 10 groups as (42-47), (48-53) ….. (98-103) and taking into consideration of mid values of each group.

Solution:

Group	N_i	Mid. Val. = x_i	x_iN_i	$(x_i - x)^2 * N_i$	$(x_i - x)^3 * N_i$
42-47	1	44,5	44,5	1242,56	-43800,3
48-53	0	50,5	0	0	0
54-59	1	56,5	56,5	540,563	-12568,1
60-65	0	62,5	0	0	0
66-71	2	68,5	137	126,563	-1423,83
72-77	8	74,5	596	27,563	-144,703
78-83	8	80,5	644	0,563	0,421875
84-89	7	86,5	605,5	45,563	307,5469
90-95	4	92,5	370	162,563	2072,672
96-101	1	98,5	98,5	351,563	6591,797
Total :			2552	2497.5	-48 964.4

$$\bar{x} = \frac{\sum x_i N_i}{N} = \frac{2552}{32} = 79.75,\ Var_x = \sum\left[(x_i - \bar{x})^2 * N_i\right] / N = 2497.5/32 = 78.05 = m_2$$

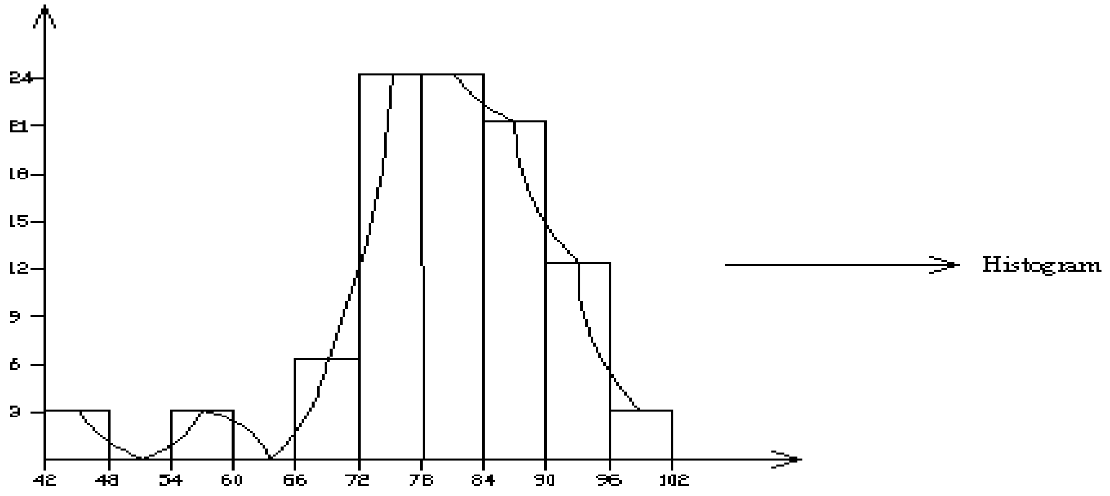
$$S_x = \sqrt{V_x} = \sqrt{78.05} = 8.83,\ m_3 = \left\{\sum\left[(x_i - \bar{x})^3 N_i\right]\right\} / N = (-48964.4)/32 = -1530.1$$

$$C_s = \frac{N^2}{(N-1)(N-2)} * \frac{m_3}{m_2^{1.5}} = \frac{32^2}{31*30} * \frac{-1530.1}{78.05^{1.5}} = -2.44 < 0 \Rightarrow \text{Left skewed distribution}$$

c. Calculate and draw the frequency histogram for the classified data.

Solution:

Group	42-47	48-53	54-59	60-65	66-71	72-77	78-83	84-89	90-95	96-101
N_i	1	0	1	0	2	8	8	7	4	1
$f_i = N_i/N)*100$	3,125	0	3,125	0	6,25	25	25	21,875	12,5	3,125
$\sum$	3,125	3,125	6,25	6,25	12,5	37,5	62,5	84,375	96,875	100



Example 3.1: The probabilities of earthquake, material failure and foundation failure of a building are 3, 6, and 8 percent, respectively. Assuming that these three events are independent:

a. Calculate the probability of (earthquake or foundation failure)

Solution:

The probabilities of the events are: P(E) = 0.03, P(M) = 0.06, P(F) = 0.08

a. Since the events of earthquake and foundation failure may occur at the same time (simultaneously), they are joint events, so, the probability of (E or F) is found by Eq. (3.5a):

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Also, since they are dependent, $P(E \cap F) = P(E) * P(F) = 0.03 * 0.08 = 0.0024$, therefore

$P(E \cup F) = P(E) + P(F) - P(E \cap F) = 0.03 + 0.08 - 0.0024 = 0.1076$ is obtained.

b. The building will be destroyed (demolished) even if one of these three events is occurred; calculate the probability of the building to be destroyed.

Solution:

There are two ways for the solution of this problem

b₁: Probability of building destroying is found by Eq. (3.5b)

$$P(E \cup M \cup F) = P(E) + P(M) + P(F) - P(E \cap M) - P(E \cap F) - P(M \cap F) + P(E \cap M \cap F) =$$
$$0.03 + 0.06 + 0.08 - 0.03 * 0.06 - 0.03 * 0.08 - 0.06 * 0.08 + 0.03 * 0.06 * 0.08 = 0.1611$$

b₂: Or, this problem can be solved as follows:

The probabilities of non occurrence of earthquake, material failure and foundation failure are:

P'(E) = 1- P(E) = 1 – 0.03 = 0.97, P'(M) = 1 – P(M) = 1 – 0.06 = 0.94, P'(F) = 1 – P(F) = 1 – 0.08 = 0.92,

In order to structure not be destroyed, none of the events becomes, in other words, all of the above non occurrence events should simultaneously become.

(Earthquake AND foundation failure AND material failure) should not become. Then the probability of non destroying is found by Eq. (3.6b) as:

$$P'(E \cap M \cap F) = P'(E) * P'(M) * P'(F) = 0.97 * 0.94 * 0.92 = 0.8389$$

Then, the probability of destroying is found as: 1 – probability of non destroying = 1 – 0.8389 = 0.1611

Example 3.2: The probabilities of flood (F), earthquake (E), and hurricane (H) in a region are 10%, 5%, and 8%, respectively. Calculate the probabilities of:

a. No event occurrence,

Solution:

a. Non occurrence probabilities are P'(F) = 1–0.10 = 0.90, P'(E) = 1–0.05 = 0.95 and P'(H) = 1– 0.08 = 0.92

Probability of no event = 0.90*0.95*0.92 = 0.7866

b. (F or E) and (F or E or H).

Solution:

These events are joint, P(A) = P(F or E) = 0.10 + 0.05 – 0.10*0.05 = 0.145

P(B) = P(F or H or H) = 1 – P(none of them) = 1 – 0.7866 = 0.2134

P(A and B) = P(A)*P(B) = 0.145*0.2134 = 0.0309

Example 3.3. The water of a city is transmitted by A, B and C pipes. The discharges of the pipes are Q_A = 10 l/s, Q_B = 15 l/s and Q_C = 25 l/s. The failure probabilities of these pipes are P_A=0.02, P_B=0.04 and P_C=0.06 and their failures are independent. Calculate the probabilities for the city:

a. Transmission of all of the water,

Solution:

There are two ways for the solution of this problem:

a_1: In order to transmission of all of the water, none of the pipes should fail. P'(A) = 0.98, P'(B) = 0.96, P'(C) = 0.94. Probability of transmission all of the water is 0.98*0.96*0.94 = 0.8844

a_2: P = 1-(0.02+0.04+0.06-0.02*0.04-0.02*0.06-0.04*0.06+0.02*0.04*0.06) = 0.8844

b. Being the transmitted discharge at least 40 l/s.

Solution:

In order to at least 40 l/s, pipe B and pipe C should not fail, P = P'(B)*P'(C) = 0.96*0.94 = 0.9024

Or P = 1 – (0.04+0.06-0.04*0.06) = 0.9024

c. Being without water.

Solution:

In order to the city being without water (no water to be transmitted) all of the three pipes should be failed. The probability is found as 0.02*0.04*0.06 = $4.8*10^{-5}$

Example 4.1: The spillway of a dam is designed according to a flood with annual probability is 0.33 percent. Calculate the non-occurring probability and 1 times and 2 times occurring probabilities of this flood in 45 years according to both Binom and Poisson Distributions.

Solution:

Binom Distribution:

p = 0.0033, q = 1 – 0.0033 = 0.9967, N=45

Eq. (4.7a): $P(x) = \frac{N!}{x!(N-x)!} p^x q^{N-x} = \frac{45!}{x!(45-x)!} * 0.0033^x * 0.9967^{45-x}$

Non-occurring = 0 times occurring x = 0 $\Rightarrow P(0) = \frac{45!}{0!(45-0)!} * 0.0033^0 * 0.9967^{45-0} = 0.8618$

x = 1 times occurring $\Rightarrow P(1) = \frac{45!}{1!(45-1)!} * 0.0033^1 * 0.9967^{45-1} = 0.1283$

x = 2 times occurring $\Rightarrow P(2) = \frac{45!}{2!(45-2)!} * 0.0033^2 * 0.9967^{45-2} = 0.00935$

Poisson Distribution:

Eq. (4.8a): $P(x) = \frac{\lambda^x}{e^{\lambda} x!}$, Eq. (4.8b): $\lambda = Np = 45 * 0.0033 = 0.1485$, $P(x) = \frac{\lambda^x}{e^{\lambda} x!} = \frac{0.1485^x}{e^{0.1485} * x!}$

Non-occurring = 0 times occurring x = 0 $\Rightarrow P(0) = \frac{0.1485^0}{e^{0.1485} * 0!} = 0.8620$

x = 1 times occurring $\Rightarrow P(1) = \frac{0.1485^1}{e^{0.1485} * 1!} = 0.1280$

x = 2 times occurring $\Rightarrow P(2) = \frac{0.1485^2}{e^{0.1485} * 2!} = 0.0095$

Example 4.2: Calculate the probability of a traffic accident occurring 0 times and 5 times in 50 years, according to Binomial and Poisson distributions, with a 20% probability to occur in any year.

Solution:

Binom Distribution: p = 0.20, q = 1 – 0.20 = 0.80

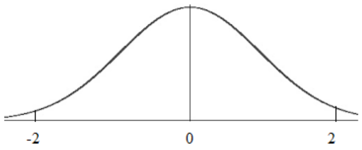
$P(0) = \frac{50!}{0!(50-0)!} 0.2^0 * 0.8^{50} = 1.427 * 10^{-5}$, $P(5) = \frac{50!}{5!(50-5)!} 0.2^5 * 0.8^{45} = 0.0295$

Poisson Distribution:

$\lambda = 50 * 0.20 = 10, \Rightarrow P(0) = \frac{0.10^1}{e^{10} * 1!} = 4.54 * 10^{-5}$,, $P(5) = \frac{10^5}{e^{10} 5!} = 0.0378$

Example 4.3: Calculate the probability of z variable is between -2 and 2.

Solution:



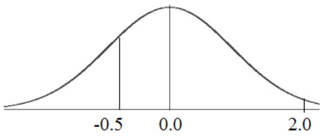
From Table 1 $\Rightarrow P(0 < z < 2) = P(-2 < z < 0) = 0.4772, \Rightarrow P(-2 < z < 2) = 0.4772 + 0.4772 = 0.9544$

Example 4.4: Calculate of area of between x = 4 and x = 9 in a Normal Distribution with a mean of 5 and standard deviation of 2. In other worlds, calculate the probability of the value is between 4 and 9.

Solution:

$\Rightarrow \bar{x} = 5, S_x = 2,$ Eq. (4.11) $\Rightarrow z = (x - \bar{x}) / S_x = (x - 5) / 2$

For $x_1 = 4 \Rightarrow z = (4 - 5)/2 = -0.5$, for $x_2 = 9 \Rightarrow z = (9 - 5)/2 = 2$



From Table 1:

$$\Rightarrow P(4 < x < 9) = P(-0.5 < z < 2) = P(-2 < z < 0) = 0.1915 + 0.4772 = 0.6687(66.87\%)$$

Example 4.5: Total annual precipitation height of a gauge station fit Normal Distribution with a mean of 90 cm and standard deviation of 25 cm. Calculate the probabilities of total annual precipitation in any year are: **a.** Less than 70 cm, **b.** Between 80 and 105 cm, **c.** Between 60 and 130 cm, **d.** More than 140 cm

Solution:

$\bar{x} = 90cm, S_x = 25cm,$ Eq. (4.11) $\Rightarrow z = (x - \bar{x}) / S_x = (x - 90)/25$

a. $P(x < 70)cm = ? \Rightarrow x = 70 \Rightarrow z = (70 - 90)/25 = -0.8,$

From Table 1 $\Rightarrow P(z < -0.8) = 0.5 - 0.2881 = 0.2129(21.19\%)$

b. $P(80 < x < 105) = ?$

$$x_1 = 80 \Rightarrow z_1 = (80 - 90)/25 = -0.4, \; x_2 = 105 \Rightarrow z_2 = (105 - 90)/25 = 0.6$$

$$P(80 < x < 105) = P(-0.4 < z < 0.6) = 0.1554 + 0.2257 = 0.3811$$

c. $P(60 < x < 130) = ?$

$$x_1 = 60 \Rightarrow z_1 = (60 - 90)/25 = -1.2, \; x_2 = 130 \Rightarrow z_2 = (130 - 90)/25 = 1.6$$

$$P(60 < x < 130) = P(-1.2 < z < 1.6) = 0.3849 + 0.4452 = 0.8301$$

d. $P(x > 140)cm = ? \Rightarrow x = 140 \Rightarrow z = (140 - 90)/25 = 2.0$,

From Table 1 $\Rightarrow P(z > 2.0) = 0.5 - 0.4772 = 0.0228(22.88\%)$

Example 4.6: Mean and standard deviation values of total annual precipitation heights of a city are 60 cm and 15 cm, respectively. Assuming that the data fit Normal Distribution:

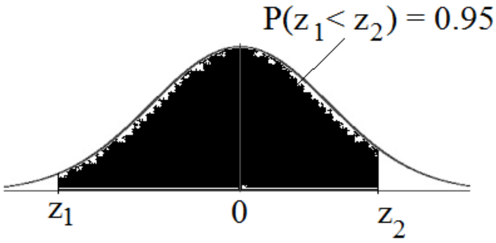
a. Determine the interval which contains 95 percent of the data (95% confidence interval),

b. A year with less than 45 cm precipitation height is called as "drought year" and with greater than 90 cm precipitation height is called as "flood year". Calculate the expected numbers of both drought and flood years in 50 years.

c. Estimate the maximum and minimum annual precipitation heights in 50 years.

Solution:

a. 95 % confidence interval means: One should find such two z_1 and z_2 values that the probability that x values is between these z values is 95 %, in other words, the area between them is 0.95



Since the z values distribute around z = 0 value symmetrically, the absolute values of z_1 and z_2 must be equal. Also, because $z_1 < 0$, $-z_1 = z_2$. Therefore, the area between $-z_1$ and 0 is equal to area between 0 and z_2.

$$P(-z_1 < z < z_2) = 0.95,\ P(-z_1 < 0) = P(z < z_2) = 0.95/2 = 0.4750$$

From Table 1, the z values corresponding to 0.4750 are read as $z_1 = -1.96$ and $z_2 = 1.96$

Eq. (4.11): $z = \dfrac{x - \bar{x}}{S_x} \Rightarrow x = \bar{x} + zS_x = 60 + 15z$

For $z_1 = -1.96 \Rightarrow x_1 = 60 + 15*(-1.96) = 30.6cm$ and

For $z_2 = 1.96 \Rightarrow x_2 = 60 + 15*1.96 = 89.4cm$

b. For a drought year, x < 45 cm should be, $x_1 = 45cm \Rightarrow z_1 = (45-60)/15 = -1$

$P(z < -1) = 0.5 - 0.3413 = 0.1587$, the expected number of drought years in 50 years is found as

$$P_i = e_i / N \Rightarrow e_i = NP_i = 50*0.1587 = 7.94 \cong 8years$$

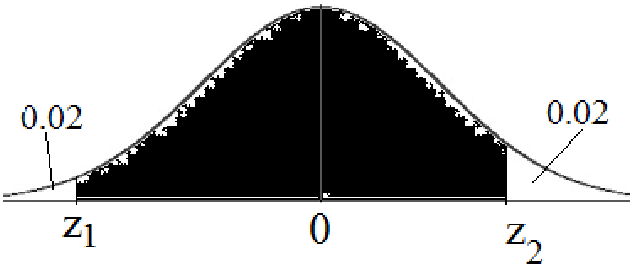
For a flood year, x > 90 cm should be, $x_2 = 90cm \Rightarrow z_2 = (90-60)/15 = 2$

$P(z > 2) = 0.5 - 0.4472 = 0.0228$, the expected number of flood years in 50 years is found as

$e_i = NP_i = 50*0.0228 = 1.14 \cong 18year$ are found.

c. For the maximum precipitation value: There is only one value equal to or greater than this value. Therefore, the probability of any value is equal to or greater than it is $P(x \geq x_{\max}) = 1/50 = 0.02$. Similarly; for the minimum precipitation value: There is only one value equal to or less than this value. Therefore, the probability of any value is equal to or less than it is $P(x \leq x_{\min}) = 1/50 = 0.02$.

The corresponding z values for $x_{\max}$ and $x_{\min}$ are found as:



$$P(z \geq z_2) = 0.02 \Rightarrow P(0 < z < z_2) = 0.5 - 0.02 = 0.48 \Rightarrow z_2 = 2.05 \Rightarrow x_{\max} = 60 + 15 * 2.05 = 90.75cm$$

$$P(z \leq z_1) = 0.02 \Rightarrow P(z_1 < z < 0) = 0.5 - 0.02 = 0.48 \Rightarrow z_1 = -2.05 \Rightarrow x_{\min} = 60 + 15 * (-2.05) = 29.25cm$$

Example 4.7: The mean annual discharge values of a stream fit Normal Distribution with a mean of 70 m^3/s and a standard deviation of 10 m^3/s. With the 80 yearly duration:

a. Estimate the number of years with less than 78 m^3/s mean discharge,

b. Calculate the 92 percent confidence interval of mean of the population.

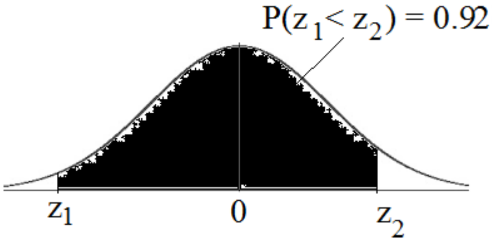
Solution:

$\bar{x} = 70m^3/s, S_x = 10m^3/s$, Eq. (4.11) $\Rightarrow z = (x - \bar{x})/S_x = (x - 70)/10$

a. $P(x < 78) = ? \Rightarrow x = 78 \Rightarrow z = (78 - 70)/10 = 0.8$,

$P(z < 0.8) = 0.5 + 0.2881 = 0.7881$, expected number of years $e_i = 80 * 0.7881 = 63 years$

b. 0.92/2 = 0.46, $z_1 = -1.75$, $z_2 = 1.75$



Eq. (4.11): $x = \bar{x} + zS_x = 70 + 10z$

For $z_1 = -1.75 \Rightarrow x_1 = 70 + 10 * (-1.75) = 52.5m^3/s$ and

For $z_1 = 1.75 \Rightarrow x_1 = 70 + 10 * 1.75) = 87.5m^3/s$ are found.

Example 4.8: Yearly maximum discharges of a stream (m^3/s) are given as follows. Calculate the probability that discharges are between (55 and 70) m^3/s, greater than 75 m^3/s; the discharge values of which exceedance probabilities are 1 and 0.1 percent; by using the distributions of: **a.** Gumbel (G), **b.** Log Normal (LN), **c.** Log Pearson Type III (LPT)

68 76 39 48 57 71 63 56 37 54 59 70 54 62 68

Solution:

Number of data N = 15, for the original values mean $\bar{x} = 58.8$ m^3/s and standard deviation S_x = 11.37 m^3/s (for Gumbel distribution), for the logarithmic values (y = lnx) $\bar{x} = 4.05$and standard deviation S_x = 0.21 (for Log Normal and Log Pearson Type III distributions).

a. Gumbel Distribution: N = 15 $\Rightarrow$From Table 2 $\Rightarrow$$y_N$ = 0.513, S_N = 1.021,

$\bar{x} = 58.8$, S_x = 11.37 m^3/s, Eq. (4.19) $\Rightarrow y_i = \dfrac{S_N}{S_x}\left[x_i - \left(\bar{x} - y_N \dfrac{S_x}{S_N}\right)\right] = \dfrac{S_N}{S_x}\left(x_i - \bar{x} + S_x \dfrac{y_N}{S_N}\right)$

For x = 55 $\Rightarrow y_i = \dfrac{1.021}{11.37}\left(55 - 58.8 + 11.37\dfrac{0.513}{1.021}\right) = 0.1718,$

Eq. (4.18a) $\Rightarrow F(x_i) = P(x \geq x_i) = 1 - e^{-e^{-y}}$, $P(x \geq 55) = 1 - e^{-e^{-0.1718}} = 0.5692$

For x = 70$\Rightarrow y_i = \dfrac{1.021}{11.37}\left(70 - 58.8 + 11.37\dfrac{0.513}{1.021}\right) = 1.5187,$ $P(x \geq 70) = 1 - e^{-e^{-1.5187}} = 0.1967,$

$P(55 < Q < 70) = 0.5692 - 0.1967 = \underline{0.3725}$

$P(55 < Q < 70) = P(Q \geq 55) - P(Q \geq 70) = 0.5692 - 0.1034 = \underline{0.4658}$

For x = 75 $\Rightarrow y_i = \dfrac{1.021}{11.37}\left(75 - 58.8 + 11.37\dfrac{0.513}{1.021}\right) = 1.9677,$ $P(x \geq 75) = 1 - e^{-e^{1.9677}} = \underline{0.1304}$

Exceedance probability 0.01, $P(x \geq x_i) = 1 - e^{-e^{-y}} = 0.01 \Rightarrow y = 4.60$

Eq. (4.19) $\Rightarrow x_i = y_i \dfrac{S_x}{S_N} + \bar{x} - S_x \dfrac{y_N}{S_N} = 4.60\dfrac{11.37}{1.021} + 58.8 - 11.37\dfrac{0.513}{1.021} = \underline{104.3 m^3 / s}$

Exceedance probability 0.001, $P(x \geq x_i) = 1 - e^{-e^{-y}} = 0.001 \Rightarrow y = 6.907$

Eq. (4.19) $\Rightarrow x_i = y_i \dfrac{S_x}{S_N} + \bar{x} - S_x \dfrac{y_N}{S_N} = 6.907\dfrac{11.37}{1.021} + 58.8 - 11.37\dfrac{0.513}{1.021} = \underline{130.0 m^3 / s}$

b. Log-Normal Distribution:

The logarithmic values are: ln(68) = 4.220, ln(76) = 4.331, …….. ln(62) = 4.127

68 76 39 48 57 71 63 56 37 54 59 70 54 62 68

4.220, 4.331, 3.664, 3.871, 4.043, 4.263, 4.143, 4.025, 3.611, 3.989, 4.078, 4.248, 3.989, 4.127, 4.220

$\bar{x} = 4.05$, S_x = 0.21, Eq (4.11) $\Rightarrow z = \dfrac{x-\bar{x}}{S_x} = \dfrac{x-4.05}{0.21}$

For Q = 55 $\Rightarrow x = \ln(55) = 4.007 \Rightarrow z = \dfrac{4.007-4.05}{0.21} = -0.20$

For Q = 70 $\Rightarrow x = \ln(70) = 4.248 \Rightarrow z = \dfrac{4.248-4.05}{0.21} = 0.95$

$P(55 < Q < 70) = P(-0.20 < z < 0.95) =$ 0.0793 + 0.3289 = $\underline{0.4082}$

For Q = 75 $\Rightarrow x = \ln(77) = 4.317 \Rightarrow z = \dfrac{4.317-4.05}{0.21} = 1.27$

$P(Q > 75) = P(z > 1.27) = 0.5 - 0.398 = \underline{0.102}$

Exceedance probability 0.01

$\Rightarrow P(z > 0.01) \Rightarrow z = 2.33 = \dfrac{x-4.05}{0.21} \Rightarrow x = 4.5393 \Rightarrow Q = e^{4.5393}$ = $\underline{93.63\ \text{m}^3/\text{s}}$

Exceedance probability 0.001

$\Rightarrow P(z > 0.001) \Rightarrow z = 3.10 = \dfrac{x-4.05}{0.21} \Rightarrow x = 4.701 \Rightarrow Q = e^{4.701}$ = $\underline{110.06\text{m}^3/\text{s}}$

c. Log Pearson Type III Distribution:

The logarithmic values are:

4.220, 4.331, 3.664, 3.871, 4.043, 4.263, 4.143, 4.025, 3.611, 3.989, 4.078, 4.248, 3.989, 4.127, 4.220

$\bar{x} = 4.05$, S_x = 0.21, Eq (3.7) $\Rightarrow z = \dfrac{x-\bar{x}}{S_x} = \dfrac{x-4.05}{0.21}$, For Q = 55 $\Rightarrow$

$x = \ln(55) = 4.007 \Rightarrow z = \dfrac{4.007-4.05}{0.21} = -0.20$,

For Q = 70 $\Rightarrow x = \ln(70) = 4.248 \Rightarrow z = \dfrac{4.248-4.05}{0.21} = 0.95$

Eq. (4.20) $x = \bar{x} + KS_x = 4.05 + 0.21K \Rightarrow K = \frac{x-4.05}{0.21}$, $K = \frac{2}{C_s}\left(1+\frac{zC_s}{6}-\frac{C_s^{\,2}}{36}\right)^3 - \frac{2}{C_s}$

Eq. (2.15.a) $\Rightarrow C_s = \frac{N^2}{(N-1)(N-2)} * \frac{m_3}{m_2^{\,1.5}}$,

m_2 is the variance $m_2 = (S_x)^2 = (0.21)^2 = 0.0441$, m_3 (the third central moment) is calculated by Eq. (2.15b)

$$m_3 = \frac{\sum(x-\bar{x})^3}{N} = \frac{(4.220-4.05)^3 - ... - (4.127-4.05)^3}{15} = -0.3114,$$

$$C_s = \frac{N^2}{(N-1)(N-2)} * \frac{m_3}{m_2^{\,1.5}} = \frac{15^2}{14*13}\frac{-0.3114}{0.441} = -0.873, \quad K = (x-4.05)/0.21,$$

$$K = \frac{2}{C_s}\left(1+\frac{zC_s}{6}-\frac{C_s^{\,2}}{36}\right)^3 - \frac{2}{C_s} = \frac{2}{-0.873}\left(1+\frac{-0.873}{6}z-\frac{(-0.873)^2}{36}\right)^3 - \frac{2}{-0873}$$

$$K = 2.291 - 2.291(0.979 - 0.1455z)^3,$$

$$Q = 55m^3/s \Rightarrow x = \ln(55) = 4.0073 \Rightarrow K = (4.0073 - 4.05)/0.21 = -0.203$$

$$K = -0.203 = 2.291 - 2.291(0.979 - 0.1455z)^3 \Rightarrow z = -0.34$$

$$Q = 70m^3/s \Rightarrow x = \ln(70) = 4.2485 \Rightarrow K = (4.2485 - 4.05)/0.21 = 0.945$$

$$K = 0.945 = 2.291 - 2.291(0.979 - 0.1455z)^3 \Rightarrow z = 0.97$$

$$P(55 < Q < 70) = P(-0.34 < z < 0.97) = 0.1331 + 0.3340 = \underline{0.4671}$$

$$Q = 75m^3/s \Rightarrow x = \ln(75) = 4.3175 \Rightarrow K = (4.3175 - 4.05)/0.21 = 1.274$$

$$K = 1.274 = 2.291 - 2.291(0.979 - 0.1455z)^3 \Rightarrow z = 1.49$$

$$P(Q > 75) = P(z > 1.49) = 0.5 - 0.4633 = \underline{0.0367}$$

Exceedance probability 0.01 $\Rightarrow P(z > (0.5 - 0.01) = 0.49 \Rightarrow z = 2.33$

$$K = 2.291 - 2.291(0.979 - 0.1455 * 2.33)^3 = 1.69$$

$$x = \bar{x} + KS_x = 4.05 + 0.21K = 4.05 + 0.21 * 1.69 = 4.4049 = \ln(Q) \Rightarrow Q = e^{4.4049} = \underline{81.9m^3/s}$$

Exceedance probability 0.001 $\Rightarrow P(z > (0.5 - 0.001) = 0.499 \Rightarrow z = 3.10$

$$K = 2.291 - 2.291(0.979 - 0.1455 * 3.10)^3 = 1.95$$

$$x = \bar{x} + KS_x = 4.05 + 0.21K = 4.05 + 0.21 * 1.95 = 4.4595 = \ln(Q) \Rightarrow Q = e^{4.4595} = \underline{86.4m^3/s}$$

Example 5.1: The mean and standard deviation values of compressive strength of 50 samples obtained from a concrete mass are 240 kg/cm^2 and 65 kg/cm^2, respectively.

a. Determine the interval which contains 95 percent of mean compressive strength of concrete,

b. Calculate the probability that the mean compressive strength is greater than 260 kg/cm^2.

Solution:

a. N = 50 > 30 $\Rightarrow$ Asymptotic distribution (Normal Distribution).

Eq. (5.2b) $\Rightarrow P_c = P(b_1 < \mu < b_2) = P\left[\left(\bar{x} - zS_x / \sqrt{N}\right) < \mu < \left(\bar{x} + zS_x / \sqrt{N}\right)\right]$

For the confidence interval with P_c probability of μ is $b_1 = \left(\bar{x} - z\frac{S_x}{\sqrt{N}}\right)$ and $b_2 = \left(\bar{x} + z\frac{S_x}{\sqrt{N}}\right)$

In these equations, z value is obtained from Normal Distribution Table (Table 4.1) corresponding the value of $(P_c / 2)$. For P$_c$ = 0.95 confidence level, P$_c$/2 = 0.475 and the corresponding z value is read as 1.96, therefore the confidence interval of the mean is calculated as:

$b_1 = \left(240 - 1.96\frac{65}{\sqrt{50}}\right) = 221.98 kg / cm^2$ and $b_2 = \left(240 + 1.96\frac{65}{\sqrt{50}}\right) = 258.02 kg / cm^2$ are obtained.

b. For sampling distribution, Eq. (4.11) $z = \frac{x - \bar{x}}{S_x}$ becomes $z = \frac{x - \bar{x}}{S_x / \sqrt{N}} = \frac{(x - \bar{x})\sqrt{N}}{S_x}$.

For x = 260 kg/cm^2, this value becomes $z = \frac{(260 - 240) * \sqrt{50}}{65} = 2.18$.

From Table 1: P > 2.18 = 0.5 – 0.4854 = 0.0146 (1.46 %) is obtained.

Example 5.2: Experiments were performed for 21 steel samples. It is found that mean and standard deviation of yielding values in the experiments are 8490 kg and 100 kg, respectively. Calculate the confidence intervals of population mean for **a.** 90 % and **b.** 98 % confidence levels.

Solution:

N = 21 < 30 $\Rightarrow$ Exact distribution (t Distribution).

Eq. (5.4) $P_c = P(b_1 < \mu < b_2) = P\left[\left(\bar{x} - tS_x / \sqrt{N-1}\right) < \mu < \left(\bar{x} + tS_x / \sqrt{N-1}\right)\right]$

Confidence interval with P_c probability of μ is $b_1 = \left(\bar{x} - t\frac{S_x}{\sqrt{N-1}}\right)$ and $b_2 = \left(\bar{x} + t\frac{S_x}{\sqrt{N-1}}\right)$

In these equations, t value is obtained from t Distribution Table (Table 5.1) by $d_f = (N-1) = 21 - 1 = 20$ degree of freedom (d$_f$) for various confidence levels.

a. P$_c$ = 0.90 and d$_f$ = 20 $\Rightarrow$ t = 1.725, therefore the confidence interval of the mean

$b_1 = \left(8490 - 1.725\frac{100}{\sqrt{20}}\right) = 8451 kg$ and $b_2 = \left(8490 + 1.725\frac{100}{\sqrt{20}}\right) = 8529 kg$ are obtained.

b. P$_c$ = 0.98 and d$_f$ = 20 $\Rightarrow$ t = 2.528, therefore the confidence interval of the mean

$b_1 = \left(8490 - 2.528\frac{100}{\sqrt{20}}\right) = 8433 kg$ and $b_2 = \left(8490 + 2.528\frac{100}{\sqrt{20}}\right) = 8547 kg$ are obtained.

Example 5.3: a. 20 and **b.** 10 samples are taken from concrete, produced a batch plant. Both groups have a 100 kg/cm^2 standard deviation. Calculate the 95% confidence interval of population standard deviation for both groups.

Solution:

Chi-Square (χ^2) distribution is used in the estimation of confidence interval of the variance. The probability of χ^2 is greater than a given value of $\chi_0{}^2$ is tabulated according to $(N-1)$ degree of freedom (d$_f$) for various confidence levels. Confidence interval of the variance is (Eq. 5.5):

$$P_c = P(b_1 < \sigma^2 < b_2) = P\left(\frac{NS^2{}_x}{\chi^2{}_{\alpha/2}} < \sigma^2 < \frac{NS^2{}_x}{\chi^2{}_{1-\alpha/2}}\right)$$

The confidence interval of the variance is calculated as:

$$b_1 = \left(\frac{NS^2{}_x}{\chi^2{}_{\alpha/2}}\right) \text{ and } b_2 = \left(\frac{NS^2{}_x}{\chi^2{}_{1-\alpha/2}}\right)$$

P$_c$ = 0.95$\Rightarrow \alpha$=1 – 0.95 = 0.05, $\Rightarrow \alpha/2 = 0.025$, $1-\alpha/2 = 0.975$

a. N = 20, d$_f$ = 20 – 1 = 19 and $\alpha/2 = 0.025$ from χ^2 Distribution Table (Table 5.2)

$\chi^2{}_{0.025} = 32.852$ and $\chi^2{}_{0.975} = 8.907$ values are read. Therefore,

for variance $\underline{b_1 = \frac{20*100^2}{32.852} = 6088 kg^2 cm^4}$, for standard deviation $\underline{b_1 = \sqrt{6088} = 78.02 kg/cm^2}$,

for variance $\underline{b_2 = \frac{20*100^2}{8.907} = 22454 kg^2 cm^4}$, for standard deviation $\underline{b_2 = \sqrt{22454} = 149.85 kg/cm^2}$ are found.

b. N = 10, d$_f$ = 10 – 1 = 9 and $\alpha/2 = 0.025$ from χ^2 Distribution Table (Table 5.2)

$\chi^2{}_{0.025} = 19.023$ and $\chi^2{}_{0.975} = 2.700$ values are read. Therefore,

for variance $\underline{b_1 = \frac{10*100^2}{19.023} = 5257 kg^2/cm^4}$, for standard deviation $\underline{b_1 = \sqrt{5257} = 72.50 kg/cm^2}$

for variance $\underline{b_2 = \frac{10*100^2}{2.700} = 37037 kg^2/cm^4}$, for standard deviation $\underline{b_1 = \sqrt{37037} = 192.45 kg/cm^2}$ are found.

Example 5.4. The mean and standard deviation values of 20 samples taken from a concrete mass are 280 kg/cm^2 and 15 kg/cm^2, respectively. Calculate the confidence interval of population of

a. mean and **b.** standard deviation for 5 % significance level.

Solution:

a. Confidence interval for the mean: N = 20 < 30 $\Rightarrow$ Exact distribution (t Distribution).

Eq. (5.4a) $P_c = P(b_1 < \mu < b_2) = P\left[\left(\bar{x} - tS_x / \sqrt{N-1}\right) < \mu < \left(\bar{x} + tS_x / \sqrt{N-1}\right)\right]$

Confidence interval with P_c probability of μ is $b_1 = \left(\bar{x} - t\frac{S_x}{\sqrt{N-1}}\right)$ and $b_2 = \left(\bar{x} + t\frac{S_x}{\sqrt{N-1}}\right)$

In these equations, t value is obtained from t Distribution Table (Table 5.1) by

$d_f = (N-1) = 20 - 1 = 19$ degree of freedom (d$_f$) for $\alpha = 0.05$ (5 %) significance level

$(P_c = 1 - 0.05 = 0.95$ confidence level) as $t = 2.093$, so, the confidence interval of the mean

$\underline{b_1 = 280 - 2.093\frac{15}{\sqrt{19}} = 272.80 kg/cm^2}$ and $\underline{b_2 = 280 + 2.093\frac{15}{\sqrt{19}} = 287.20 kg/cm^2}$ is obtained.

b. Confidence interval for variance is (Eq. 5.5): $b_1 = \left(\frac{NS^2{}_x}{\chi^2{}_{\alpha/2}}\right)$ and $b_2 = \left(\frac{NS^2{}_x}{\chi^2{}_{1-\alpha/2}}\right)$

P$_c$ = 0.95 $\Rightarrow \alpha$ = 1 – 0.95 = 0.05, $\Rightarrow \alpha/2 = 0.025$, $1 - \alpha/2 = 0.975$

from χ^2 Distribution Table (Table 5.2) $\chi^2{}_{0.025} = 32.852$ and $\chi^2{}_{0.975} = 8.907$ values are read. Therefore,

for variance $\underline{b_1 = \frac{NS_x{}^2}{\chi^2{}_{\alpha/2}} = \frac{20*15^2}{32.852} = 136.98 kg^2/cm^4}$, $\underline{b_2 = \frac{NS_x{}^2}{\chi^2{}_{1-\alpha/2}} = \frac{20*15^2}{8.907} = 505.22 kg^2/cm^4}$,

for standard deviation $\underline{b_1 = \sqrt{136.98} = 11.70 kg/cm^2}$, $\underline{b_2 = \sqrt{505.22} = 22.48 kg/cm^2}$ are found.

Example 6.1. Lighting is applied to decrease traffic accidents in a junction. Accident numbers are given below before lighting for 12 months (x_1) and after lighting for 8 months (x_2). Decide whether or not the lighting has affected **a.** the mean and **b.** standard deviation of traffic accident for $\alpha = 0.01$.

x_1	5	8	11	7	6	9	6	8	6	7	7	10
x_2	3	6	7	5	11	8	7	2	-	-	-	-

Solution:

For the given data, the mean and standard deviations are calculated as follows:

$N_1 = 12,\ \bar{x}_1 = 7.5,\ S_1 = 1.784,\ N_2 = 8,\ \bar{x}_2 = 6.125,\ S_2 = 2.85$

a. The null hypothesis: H_0: $\mu_1 = \mu_2$: The means of two different populations are the same, in other words, the lighting HAS NOT (CHANGED) AFFECTED the mean of the accidents,

The alternative hypothesis: H_1: $\mu_1 \neq \mu_2$: The means of two different populations are NOT the same, in other words, the lighting HAS AFFECTED the mean of the accidents.

$\alpha = 0.01 \Rightarrow 0.5 - \alpha/2 = 0.495 \Rightarrow$ Table 4.1: $z = 2.58$

Eq. (6.1) $\Rightarrow S_{\bar{x}} = \sqrt{\dfrac{(N_1 S_2^{\ 2} + N_2 S_1^{\ 2})}{N_1 N_2}} = \sqrt{\dfrac{12*2.85^2 + 8*1.784^2}{12*8}} = 1.132,\ \Delta\bar{x} = \bar{x}_1 - \bar{x}_2 = 7.5 - 6.125 = 1.375$

$|\Delta\bar{x}| = 1.375,\ zS_{\bar{x}} = 2.58*1.132 = 2.921$, if $|\Delta\bar{x}| \leq zS_{\bar{x}}$, then the H_0 is accepted. Since $|\Delta\bar{x}| < zS_{\bar{x}} \Rightarrow$ H_0: $\mu_1 = \mu_2$ hypothesis is accepted for $\alpha = 0.01$ significance level (1 - 0.01 = 0.99 confidence level).

THE LIGHTING HAS NOT AFFECTED (CHANGED) THE MEAN OF THE ACCIDENTS.

b. The null hypothesis: H_0: $\sigma_1 = \sigma_2$: The standard deviations of two different populations are the same, in other words, the lighting HAS NOT AFFECTED the standard deviation of the accidents,

The alternative hypothesis: $\sigma_1 \neq \sigma_2$: The standard deviations of two different populations are the same, in other words, the lighting HAS AFFECTED (CHANGED) the standard deviation of the accidents.

Eq. (6.2) $\Rightarrow$ If $S_1 > S_2 \Rightarrow F_c = \dfrac{S_1^{\ 2}}{S_2^{\ 2}}$, if $S_2 > S_1 \Rightarrow F_c = \dfrac{S_2^{\ 2}}{S_1^{\ 2}}$

$$S_1 = 1.784 < S_2 = 2.85 \Rightarrow F_c = \frac{S_2^{\ 2}}{S_1^{\ 2}} = \frac{2.85^2}{1.784^2} = 2.552$$

$\alpha = 0.01$, $m = N_1 - 1 = 12 - 1 = 11$ and $n = N_2 - 1 = 8 - 1 = 7$ $\Rightarrow$ F Distribution (Table 6.1) by linear interpolation $F_t = 6.56$ is read. If the computed F_c is less than its tabulated value (F_t) then H_0 hypothesis is accepted. Since $F_c < F_t \Rightarrow H_0$: $\sigma_1 = \sigma_2$ hypothesis is accepted for $\alpha = 0.01$ significance level (1-0.01=0.99 confidence level). THE LIGHTING HAS NOT AFFECTED (CHANGED) THE STANDARD DEVIATION OF THE ACCIDENTS.

Example 6.2: The number (N), mean ($\bar{x}$) and standard deviations (S_x) of precipitation height values before and after dam construction are given as follows. Determine whether or not dam construction has changed **a.** the mean and **b.** the standard deviation of precipitation heights for α=0.01.

	N	$\bar{x}$ (cm)	S_x (cm)
Before the dam	20	100	23
After the dam	25	116	10

Solution:

a. The null hypothesis: H_0: $\mu_1 = \mu_2$: The means of two different populations are the same, in other words, the dam construction has not changed the mean of precipitation heights,

$\alpha = 0.01 \Rightarrow 0.5 - \alpha/2 = 0.495 \Rightarrow$ Table 4.1: z = 2.58

Eq. (6.1) $\Rightarrow S_{\bar{x}} = \sqrt{\frac{(N_1 S_2{}^2 + N_2 S_1{}^2)}{N_1 N_2}} = \sqrt{\frac{20*10^2 + 25*23^2}{20*25}} = 5.52$, $\Delta\bar{x} = \bar{x}_1 - \bar{x}_2 = 100 - 116 = -16$

$|\Delta\bar{x}| = 16$, $zS_{\bar{x}} = 2.58*5.52 = 14.24$, since $|\Delta\bar{x}| > zS_{\bar{x}} \Rightarrow$ H_0: $\mu_1 = \mu_2$ hypothesis is rejected for α = 0.01 significance level.

THE DAM CONSTRUCTION HAS CHANGED THE MEAN OF PRECIPITATION HEIGHTS,

b. The null hypothesis: H_0: $\sigma_1 = \sigma_2$: The standard deviations of two different populations are the same, in other words, the lighting has not changed the standard deviation of the accidents,

Eq. (6.2) $\Rightarrow$ If $S_1 > S_2 \Rightarrow F_c = \frac{S_1{}^2}{S_2{}^2}$, if $S_2 > S_1 \Rightarrow F_c = \frac{S_2{}^2}{S_1{}^2}$

$$S_1 = 23 > S_2 = 10 \Rightarrow F_c = \frac{S_1{}^2}{S_2{}^2} = \frac{23^2}{10^2} = 5.29$$

$\alpha = 0.01$, $m = N_1 - 1 = 20 - 1 = 19$ and $n = N_2 - 1 = 25 - 1 = 24$ $\Rightarrow$ F Distribution (Table 6.1) by linear interpolation $F_t = 2.77$ is read. Since $F_c > F_t \Rightarrow H_0$: $\sigma_1 = \sigma_2$ hypothesis is rejected for α = 0.01 significance level

THE DAM CONSTRUCTION HAS CHANGED THE STANDARD DEVIATION OF PRECIPITATION HEIGHTS.

Example 6.3: Numbers of occurring (x) and observation (O) of a flood, of which probability is 5 percent, are given as follows, for a 40 yearly observation period. Determine whether the data fit Binom and Poisson Distributions for 5% significance level.

Solution:

Number of data N = 40, number of groups: m = 6 (0, 1, 2, 3, 4 and 5),

probability of success p = 5 % = 0.05, probability of fail q = 1 – p = 1 – 0.05 = 0.95,

a. CALCULATION OF PROBABILITIES AND EXPECTED VALUES:

Binom Distribution: Eq. (4.7a) $\Rightarrow P(x) = p_x = \frac{N!}{x!(N-x)!} p^x q^{N-x} = \frac{40!}{x!(40-x)!} 0.05^x 0.95^{40-x}, \; 0 \le x \le 5$

Expected value $e_x = p_x * N = 40 * p_x$. For example:

x = 0 $\Rightarrow P(0) = p_0 = \frac{40!}{0!(40-0)!} 0.05^0 0.95^{40-0} = 0.95^{40} = 0.1285$, e_0 = 40*0.1285 = 5.14

x = 1 $\Rightarrow P(1) = p_1 = \frac{40!}{1!(40-1)!} 0.05^1 0.95^{39} = 0.2706$, e_1 = 40*0.2706 = 10.822

Similar calculations are made and presented in the following table.

Poisson Distribution:

Eq. (4.8a) $\Rightarrow P(X = x) = p_x = \frac{\lambda^x}{e^\lambda x!}$, $\lambda = Np = 40 * 0.05 = 2.0$, $\Rightarrow p_x = \frac{2^x}{e^2 x!}$, $0 \le x \le 5$

For example: x = 0 $\Rightarrow P(0) = p_0 = \frac{2^0}{e^2 0!} = 0.1353$, e_0 = 40*0.1353 = 5.413

x = 1 $\Rightarrow P(1) = p_1 = \frac{2^1}{e^2 1!} = 0.2707$, e_1 = 40*0.2707 = 10.827

GIVEN DATA							
x_i		0	1	2	3	4	5
O_i		6	10	9	8	5	2
SOLUTION							
BINOM	P_i	0.1353	0.2706	0.2777	0.1851	0.0901	0.0342
	e_i	5.140	10.822	11.107	7.405	3.605	1.366
POISSON	P_i	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361
	e_i	5.413	10.827	10.827	7.218	3.609	1.444

b. CALCULATION OF CHI-SUARE VALUE:

Eq. (6.3) $\chi^2{}_c \Rightarrow = \sum_{i=1}^{m}\left[\frac{(O_i - e_i)^2}{e_i}\right]$

Binom Distribution: $\chi^2{}_c = \frac{\left(6-5.14)^2\right)}{5.14} + \frac{(10-10.822)^2}{10.822} + ... + \frac{(2-1.366)^2}{1.366} = 1.488$

Poisson Distribution: $\chi^2{}_c = \frac{\left(6-5.413)^2\right)}{5.413} + \frac{(10-10.827)^2}{10.827} + ... + \frac{(2-1.444)^2}{1.444} = 1.27$

c. DECISION:

If $\chi^2{}_c \le \chi^2{}_t$ then the sample fits the theoretical distribution, and vice versa. χ^2_t is read for 1-α . In reading the tabulated χ^2 values, the degree of freedom is m-2 for Poisson and m-3 for Binom distribution.

Binom Distribution: Degree of freedom m–3 = 6–3 = 3 and 5% significance level (α = 0.05), 1 - α = 0.95

From Table 5.2$\Rightarrow \chi^2{}_t = 7.815$ is read.

Conclusion: Since $\chi^2{}_c = 1.488 < \chi^2{}_t = 7.815 \Rightarrow$ The data fit Binom Distribution for 5% significance level.

Poisson Distribution: Degree of freedom m–2 = 6–2 = 4 and 5% significance level (α = 0.05), 1-α = 0.95

From Table 5.2$\Rightarrow \chi^2{}_t = 9.488$ is read.

Conclusion: Since $\chi^2{}_c = 1.27 < \chi^2{}_t = 9.488 \Rightarrow$ The data fit Poisson Distribution for 5% significance level.

Example 6.4: The 80 yearly annual total precipitation height values of a region fit Normal Distribution with a mean of 70 cm and a standard deviation of 10 cm. It is aimed to test whether or not the data fit Normal Distribution. There are 50 data between (60-75) cm, calculate the chi-square (χ^2) value.

Solution:

$\bar{x} = 70, S_x = 10$, N = 80, O_i = 50,. If the data fits Normal Distribution, then, the probability of the data is between (60-75) cm is calculated as was given in Normal Distribution:

$z = \frac{x-\bar{x}}{S_x} = \frac{x-70}{10}$, for 60 cm $\Rightarrow z = \frac{60-70}{10} = -1$, for 75 cm $\Rightarrow z = \frac{75-70}{10} = 0.5$

$P(60 < x < 75) = P(-1 < z < 0.5) = 0.3413 + 0.1915 = 0.5328$

Expected value $e_i = N*p_i = 80*0.5328 = 42.624$, observed value $o_i = 50$

$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i - e_i)^2}{e_i}\right] = \frac{\left(50-42.624)^2\right)}{42.624} = 1.276$ is obtained.

Example 6.5: The distribution of marks taken by 80 students in an exam is given as follows. The mean is 50 and the standard deviation is 16. Determine whether or not the marks fit normal distribution for both 1% and 5% significance levels.

GROUP	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
NO OF STUD. (O$_i$)	1	5	8	10	14	15	11	8	6	2

Solution:

Number of data N = 80, number of groups m = 10, $\bar{x} = 50, S_x = 16$, N = 80, $z = \frac{x-\bar{x}}{S_x} = \frac{x-50}{16}$,

For example, for the first group (0-10), the values are 0 and 10 and the corresponding z values are calculated as follows:

If the data fits Normal Distribution, then, the probability of the data is between (0 - 10) is calculated as was given in Normal Distribution:

$$x_1 = 0 \Rightarrow z_1 = \frac{0-50}{16} = -3.13,\ x_2 = 10 \Rightarrow z_2 = \frac{10-50}{16} = -2.50.$$

$$P(0 < x < 10) = P(-3.13 < z < -2.50) = 0.4991 - 0.4938 = 0.0053$$

Expected value e$_i$ = N*p$_i$ = 80*0.0053 = 0.424, observed value (no of student) O$_i$ = 1

$$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i - e_i)^2}{e_i}\right] = \frac{\left(1-0.424)^2\right)}{0.424} = 0.782 \text{ is obtained.}$$

For the second group (10-20), values are 10 and 20 and the corresponding z values are calculated as follows:

If the data fits Normal Distribution, then, the probability of the data is between (10 - 20) is calculated as was given in Normal Distribution:

$$x_1 = 10 \Rightarrow z_1 = \frac{10-50}{16} = -2.50,\ x_2 = 20 \Rightarrow z_2 = \frac{20-50}{16} = -1.88$$

$$P(10 < x < 20) = P(-2.50 < z < -1.88) = 0.4938 - 0.4699 = 0.0239$$

Expected value e$_i$ = N*p$_i$ = 80*0.0239 = 1.912, observed value (no of student) O$_i$ = 5

$$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i - e_i)^2}{e_i}\right] = \frac{(5-1.912)^2}{1.912} = 4.987 \text{ is obtained.}$$

Similar calculations are made for all of the groups and the results are presented in the table:

GROUP	0-10	10-20	20-30	30-40	40-50
NO OF ST. (O_i)	1	5	8	10	14
x_1	0	10	20	30	40
x_2	10	20	30	40	50
z_1	-3.13	-2.50	-1.88	-1.25	-0.63
z_2	-2.50	-1.88	-1.25	-0.63	0.00
P_i	0.0053	0.0239	0.0755	0.1587	0.2357
$e_i=80*P_i$	0.424	1.912	6.040	12.696	18.856
$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i-e_i)^2}{e_i}\right]$	0.782	4.987	0.636	0.572	1.251

GROUP	50-60	60-70	70-80	80-90	90-100
NO OF ST. (O_i)	15	11	8	6	2
x_1	50	60	70	80	90
x_2	60	70	80	90	100
z_1	0.00	0.63	1.25	1.88	2.50
z_2	0.63	1.25	1.88	2.50	3.13
P_i	0.2357	0.1587	0.0755	0.0239	0.0053
$e_i = 80*P_i$	18.856	12.696	6.040	1.912	0.424
$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i-e_i)^2}{e_i}\right]$	0.789	0.227	0.636	8.740	5.858

$\chi^2{}_c = \sum_{i=1}^{m}\left[\frac{(O_i-e_i)^2}{e_i}\right] = 0.782 + 4.987 + ... + 8.740 + 5.858 = 24.478$ is calculated.

If $\chi^2{}_c \le \chi^2{}_t$ then the sample fits the theoretical distribution, and vice versa. $\chi^2{}_t$ is read for 1-α . In reading the tabulated χ^2 values, the degree of freedom is d_f = m-3 for Normal distribution. d_f = m – 3 = 10 – 3 =7 for 1% significance level (α = 0.01), 1 - α = 0.99, From Table 5.2 $\Rightarrow \chi^2{}_t = 18.475$ is read.

Conclusion: Since $\chi^2{}_c = 24.878 > \chi^2{}_t = 18.475 \Rightarrow$ The data do not fit Normal Distribution for 5% significance level.

for 5% significance level (α = 0.05), 1 - α = 0.95, From Table 5.2 $\Rightarrow \chi^2{}_t = 14.067$ is read.

Conclusion: Since $\chi^2{}_c = 24.878 > \chi^2{}_t = 14.0675 \Rightarrow$ The data do not fit Normal Distribution for 1% significance level.