1)

a) $\bar{x} = 7\ mm$
$s = 2.64\ mm$
$n = 20$

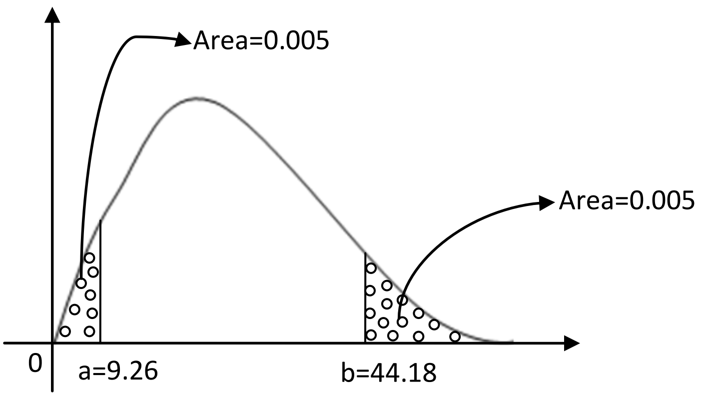
Use Students' t-distribution as σ is not known and n is less than 30.

$$P\left(-a \leq \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} \leq a\right) = 0.98$$

From Students' t-distribution table, a=2.500 for P=0.01 and df=n-1=24-1=23

$$-2.500 \leq \frac{7-\mu}{\frac{2.64}{\sqrt{24}}} \leq 2.500$$
$$5.653 \leq \mu \leq 8.347$$

b) $P\left(a \leq \frac{(n-1)*s^2}{\sigma^2} \leq b\right) = 0.99$



df=n-1=24-1=23

$$9.26 \leq \frac{(24-1)*2.64^2}{\sigma^2} \leq 44.18$$
$$1.905 \leq \sigma \leq 17.311$$

c) The population mean cannot be 10 mm with a significance level of 2% as the interval for population mean with a significance level of 2% is [5.653, 8.347].

2)

a) $\bar{x} = 672\ cm$
$\sigma = 15\ cm$

Use Normal Distribution as σ is known.

$$P\left(-a \leq \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq a\right) = 0.95$$

From normal distribution table, a=1.96 for P=0.975

$$-1.96 \leq \frac{672-\mu}{\frac{15}{\sqrt{50}}} \leq 1.96$$
$$667.84 \leq \mu \leq 676.16$$

The length of the confidence interval is

L=676.16-667.84=8.32

a) $P(|\bar{x}-\mu| \leq 5) = P(-5 \leq \bar{x}-\mu \leq 5) = P\left(\frac{-5}{15/\sqrt{n}} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq \frac{5}{15/\sqrt{n}}\right) = 0.95$

$$\frac{5}{\frac{15}{\sqrt{n}}} = 1.96 \;\Rightarrow n = 34.6$$

Therefore, take n=35.

3)

a)

i) The probabilities for the given x values are calculated by using the Binomial distribution.

$$P(x=0) = \binom{5}{0} * \theta^0 * (1-\theta)^{5-0} = 1 * 0.10^0 * (1-0.10)^5 = 0.591$$
$$P(x=1) = \binom{5}{1} * \theta^1 * (1-\theta)^{5-1} = 5 * 0.10^1 * (1-0.10)^4 = 0.328$$
$$P(x=2) = \binom{5}{2} * \theta^2 * (1-\theta)^{5-2} = 10 * 0.10^2 * (1-0.10)^3 = 0.073$$
$$P(x=3) = \binom{5}{3} * \theta^3 * (1-\theta)^{5-3} = 10 * 0.10^3 * (1-0.10)^2 = 0.008$$
$$P(x=4) = \binom{5}{4} * \theta^4 * (1-\theta)^{5-4} = 5 * 0.10^4 * (1-0.10)^1 = 4.5 * 10^{-4}$$
$$P(x=5) = \binom{5}{5} * \theta^5 * (1-\theta)^{5-4} = 1 * 0.10^5 * (1-0.10)^0 = 1 * 10^{-5}$$

The goodness of fit test is

X	**o_i**	**Probability**	**e_i**	**$(o_i-e_i)^2/e_i$**
0	2	0.591	2.955	0.309
1	1	0.328	1.640	0.250
2	1	0.073	0.365	1.105
3	1	0.008	0.040	23.040
4	0	$4.5*10^{-4}$	0.002	0.002
5	0	$1*10^{-5}$	0.000	0.000
				24.705

n	5
k	6
r	0

$$df = k - r - 1 = 6 - 0 - 1 = 5$$
$$\chi_{0.05,5} = 11.07 < \sum \frac{(o_i - e_i)^2}{e_i} = 24.705$$

No Good

Therefore, the binomial distribution assumption is not reasonable.

ii) The Kolmogorov-Smirnov Test is summarized in the below table.

Occurence	Observed Frequency	Probability	Theoretical Cumulative Relative Probability ($F_x(x)$)	Observed Relative Probability	Observed Cumulative Relative Probability ($S_x(x)$)	$\|F_x(x)-S_x(x)\|$
0	2	0.591	0.591	0.400	0.400	0.191
1	1	0.328	0.919	0.200	0.600	0.319
2	1	0.073	0.992	0.200	0.800	0.192
3	1	0.008	1.000	0.200	1.000	0.000
4	0	0.00045	1.000	0.000	1.000	0.000
5	0	0.00001	1.000	0.000	1.000	0.000
	5				Maximum	0.319

From Kolmogorov-Smirnov table, the acceptable limit is 0.56 for a degree of freedom of 5. (dof=n)

$$|F_x(x)-S_x(x)|_{max} = 0.319 < 0.56$$

Check

Therefore, the binomial distribution assumption is reasonable.

b)

i) This time, θ should be found from the above data set.

$$\bar{x} = \frac{\sum_{x=0}^{5} x * f_x(x)}{n} = \frac{0*2+1*1+2*1+3*1+4*0+5*0}{5} = \frac{6}{5} = 1.2$$

$$E(x) = n*\theta \;\Rightarrow \theta = \frac{E(x)}{n} = \frac{1.2}{5} = 0.24$$

The probabilities for the given x values are calculated by using the Binomial distribution.

$$P(x=0) = \binom{5}{0} * \theta^0 * (1-\theta)^{5-0} = 1*0.24^0*(1-0.24)^5 = 0.254$$
$$P(x=1) = \binom{5}{1} * \theta^1 * (1-\theta)^{5-1} = 5*0.24^1*(1-0.24)^4 = 0.400$$
$$P(x=2) = \binom{5}{2} * \theta^2 * (1-\theta)^{5-2} = 10*0.24^2*(1-0.24)^3 = 0.253$$
$$P(x=3) = \binom{5}{3} * \theta^3 * (1-\theta)^{5-3} = 10*0.24^3*(1-0.24)^2 = 0.08$$
$$P(x=4) = \binom{5}{4} * \theta^4 * (1-\theta)^{5-4} = 5*0.24^4*(1-0.24)^1 = 0.013$$
$$P(x=5) = \binom{5}{5} * \theta^5 * (1-\theta)^{5-4} = 1*0.24^5*(1-0.24)^0 = 7.96*10^{-4}$$

The goodness of fit test is

X	o_i	Probability	e_i	$(o_i-e_i)^2/e_i$
0	2	0.254	1.270	0.420
1	1	0.400	2.000	0.500
2	1	0.253	1.265	0.056

X	o_i	Probability	e_i	$(o_i-e_i)^2/e_i$
3	1	0.08	0.400	0.900
4	0	0.013	0.065	0.065
5	0	$7.96*10^{-4}$	0.004	0.004
				1.944

n	5	
k	6	
θ	0.24	r=1

$$df = k - r - 1 = 6 - 1 - 1 = 4$$

$$\chi_{0.05,4} = 9.49 > \sum \frac{(o_i - e_i)^2}{e_i} = 24.705$$

Check

Therefore, the binomial distribution assumption is reasonable.

ii) The Kolmogorov-Smirnov Test is summarized in the below table.

Occurence	Observed Frequency	Probability	Theoretical Cumulative Relative Probability ($F_x(x)$)	Observed Relative Probability	Observed Cumulative Relative Probability ($S_x(x)$)	$\lvert F_x(x)-S_x(x)\rvert$
0	2	0.254	0.254	0.400	0.400	0.146
1	1	0.4	0.654	0.200	0.600	0.054
2	1	0.253	0.907	0.200	0.800	0.107
3	1	0.08	0.987	0.200	1.000	0.013
4	0	0.013	1.000	0.000	1.000	0.000
5	0	0.000796	1.001	0.000	1.000	0.001
	5				Maximum	0.146

From Kolmogorov-Smirnov table, the acceptable limit is 0.56 for a degree of freedom of 5. (dof=n)

$$|F_x(x)-S_x(x)|_{max} = 0.146 < 0.56$$

Check

Therefore, the binomial distribution assumption is reasonable.

4)

a) The sample mean and the sample standard deviations are calculated as follows:

Class Interval (MPa)	Class Mark (x)	Frequency (f)	x*f	$(x-\mu)^2*f$
90-95	92.5	4	370	524.41
95-100	97.5	19	1852.5	790.4475
100-105	102.5	39	3997.5	81.9975
105-110	107.5	23	2472.5	289.8575

Class Interval (MPa)	Class Mark (x)	Frequency (f)	x*f	(x-μ)^2*f
110-115	112.5	12	1350	877.23
115-120	117.5	3	352.5	550.8075
		100	10395	3114.75

$$\bar{x} = \frac{\sum_{i=1}^{100} x_i * f_i}{n} = \frac{10395}{100} = 103.95\ MPa$$

$$s = \sqrt{\frac{\sum_{i=1}^{100}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{3114.75}{100-1}} = 5.609\ MPa$$

b) $\left.\begin{array}{l} \bar{x} = 103.95\ MPa \\ s = 5.609\ MPa \\ n = 100 \end{array}\right\}$ Use Normal Distribution as n is greater than 30.

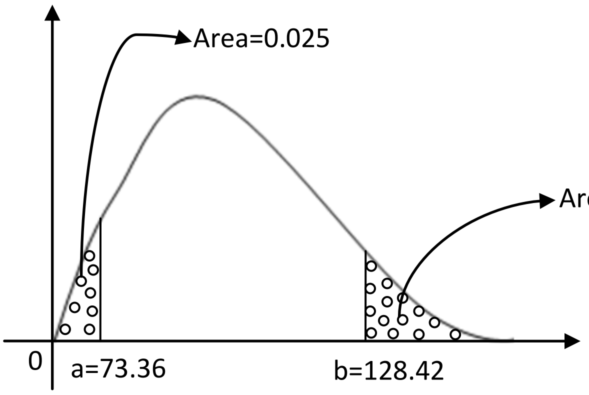
$$P\left(-a \le \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \le a\right) = 0.99$$

From normal distribution table, a=2.575 for P=0.995

$$-2.575 \le \frac{103.95 - \mu}{\frac{5.609}{\sqrt{100}}} \le 2.575$$

$$102.51 \le \mu \le 105.39$$

c) $P\left(a \le \frac{(n-1)*s^2}{\sigma^2} \le b\right) = 0.95$



df=n-1=100-1=99

$$73.36 \le \frac{(100-1)*5.609^2}{\sigma^2} \le 128.42$$

$$4.925 \le \sigma \le 6.516$$

d)

i) The probabilities for the given x values are calculated by using the normal distribution.

$$P(90 < x < 95) = P\left(\frac{90 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{95 - 103.95}{5.609}\right) = P(-2.487 < z < -1.596)$$

$$= P(z < 2.487) - P(z < 1.596) = 0.9935 - 0.9448 = 0.0487$$

$$P(95 < x < 100) = P\left(\frac{95 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{100 - 103.95}{5.609}\right) = P(-1.596 < z < -0.704)$$
$$= P(z < 1.596) - P(z < 0.704) = 0.9448 - 0.7592 = 0.1856$$

$$P(100 < x < 105) = P\left(\frac{100 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{105 - 103.95}{5.609}\right) = P(-0.704 < z < 0.187)$$
$$= P(z < 0.187) - \big(1 - P(z < 0.704)\big) = 0.5741 - (1 - 0.7592) = 0.3333$$

$$P(105 < x < 110) = P\left(\frac{105 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{110 - 103.95}{5.609}\right) = P(0.187 < z < 1.079)$$
$$= P(z < 1.079) - P(z < 0.187) = 0.8597 - 0.5741 = 0.2856$$

$$P(110 < x < 115) = P\left(\frac{110 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{115 - 103.95}{5.609}\right) = P(1.079 < z < 1.97)$$
$$= P(z < 1.97) - P(z < 1.079) = 0.9756 - 0.8597 = 0.1159$$

$$P(115 < x < 120) = P\left(\frac{115 - 103.95}{5.609} < \frac{x - \mu}{\sigma} < \frac{120 - 103.95}{5.609}\right) = P(1.97 < z < 2.861)$$
$$= P(z < 2.861) - P(z < 1.97) = 0.9979 - 0.9756 = 0.0223$$

The goodness of fit test is

Class Interval (MPa)	Class Mark (x)	o_i	Probability	e_i	$(o_i-e_i)^2/e_i$
90-95	92.5	4	0.0487	4.870	0.155
95-100	97.5	19	0.1856	18.560	0.010
100-105	102.5	39	0.3333	33.330	0.965
105-110	107.5	23	0.2856	28.560	1.082
110-115	112.5	12	0.1159	11.590	0.015
115-120	117.5	3	0.0223	2.230	0.266
					2.493

n	5	
k	6	
μ	103.95	r=2
σ	5.609	

$$df = k - r - 1 = 6 - 2 - 1 = 3$$
$$\chi_{0.05,3} = 7.81 > \sum \frac{(o_i - e_i)^2}{e_i} = 2.493$$

Check

Therefore, the normal distribution assumption is reasonable.

i) The Kolmogorov-Smirnov Test is summarized in the below table.

Occurence	Observed Frequency	Probability	Theoretical Cumulative Relative Probability ($F_x(x)$)	Observed Relative Probability	Observed Cumulative Relative Probability ($S_x(x)$)	$\lvert F_x(x)-S_x(x)\rvert$
92.5	4	0.0487	0.049	0.040	0.040	0.009
97.5	19	0.1856	0.234	0.190	0.230	0.004
102.5	39	0.3333	0.568	0.390	0.620	0.052

Occurence	Observed Frequency	Probability	Theoretical Cumulative Relative Probability ($F_x(x)$)	Observed Relative Probability	Observed Cumulative Relative Probability ($S_x(x)$)	$\|F_x(x)-S_x(x)\|$
107.5	23	0.2856	0.853	0.230	0.850	0.003
112.5	12	0.1159	0.969	0.120	0.970	0.001
117.5	3	0.0223	0.991	0.030	1.000	0.009
	100				Maximum	0.052

From Kolmogorov-Smirnov table, the acceptable limit is 0.136 for a degree of freedom of 100. (dof=n)

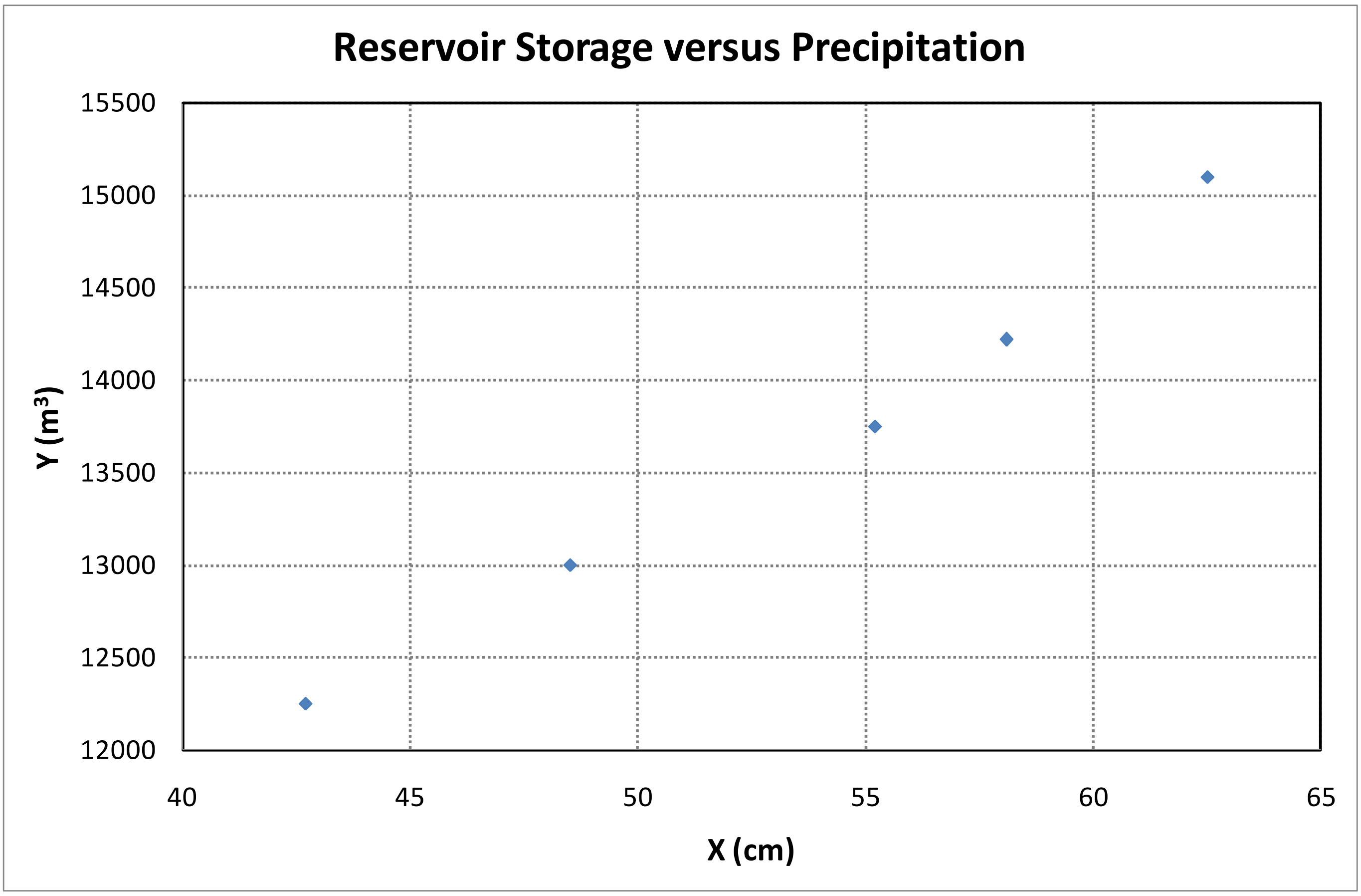
$$|F_x(x)-S_x(x)|_{max} = 0.052 < 0.136$$

Check

Therefore, the normal distribution assumption is reasonable.

5)

a) The scatter diagram of the given data set is shown below.



It can easily be inferred from the above graph that there is a strong linear relation between the reservoir storage and the precipitation.

x	y	$x-x_{mean}$	$y-y_{mean}$	$(x-x_{mean})^2$	$(y-y_{mean})^2$	$(x-x_{mean})(y-y_{mean})$	$\hat{y}$	$(\hat{y}-y)^2$	$(\hat{y}-y_{mean})^2$
62.5	15100	9.100	1436	82.810	2062096	13067.6	14926.540	30088.323	1594007.604
42.7	12250	-10.700	-1414	114.490	1999396	15129.8	12179.475	4973.807	2203815.126
48.5	13000	-4.900	-664	24.010	440896	3253.6	12984.171	250.567	462167.8851

x	y	$x-x_{mean}$	$y-y_{mean}$	$(x-x_{mean})^2$	$(y-y_{mean})^2$	$(x-x_{mean})(y-y_{mean})$	$\hat{y}$	$(\hat{y}-y)^2$	$(\hat{y}-y_{mean})^2$
55.2	13750	1.800	86	3.240	7396	154.8	13913.733	26808.565	62366.67838
58.1	14220	4.700	556	22.090	309136	2613.2	14316.081	9231.591	425209.8535
				246.640	4818920	34219	68320	71352.854	4747567.146

x_{mean}	53.400
y_{mean}	13664

SS_{xx}	246.640
SS_{yy}	4818920
SS_{xy}	34219

β	138.741
α	6255.248

The regression line equation is

$$y = 138.71 * x + 6255.248$$

b) $R^2 = 1 - \frac{SSE}{SS_{yy}} = 1 - \frac{71352.854}{4818920} = 0.985 \Rightarrow R = 0.993$

Since the correlation coefficient is very close to 1, there is a strong linear relation between x and y. Moreover, 98.5% of y values are well represented by x values as the coefficient of determination is 0.985.

c) $b = \beta \pm t_{\alpha/2,n-2}\frac{S_{y/x}}{\sqrt{SS_{xx}}} = \beta \pm t_{0.005,3}\frac{S_{y/x}}{\sqrt{SS_{xx}}} = 138.71 \pm 5.841 * \frac{154.222}{\sqrt{246.64}} = 138.71 \pm 57.359$

d) $y_{new} = \hat{y} \pm t_{\alpha/2,n-2}S_{y/x}\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{SS_{xx}} + 1}$

$$= 13192.282 \pm 3.182 * 154.222 * \sqrt{\frac{1}{5} + \frac{(50-53.4)^2}{246.64} + 1}$$

$$= 13192.282 \pm 547.97$$