- Start every new question from a new page
- Use the numbering for the subparts as mentioned in the question, or else the examiner wouldn't check that particular question and the question gets disqualified
- Use left alignment throughout if possible, looks more preofessional
- So presentation wise, the submission is not at all upto mark, need to work over this

Question 1

Given,

$$X_1, X_2 \dots X_n \sim N(\mu, \sigma^2)$$

1) The shape of sampling distribution will be positively skewed in starting as n tends toward ∞, the shape of distribution will be symmetric.
limited explanation provided for a 2 marker, first you were required to define what is the distribution, then only comment on the shape.
Also do not write at the starting, use technical word like small sample size n and large sample size n
2) As we know,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$applying\ variance\ both\ sides,$

$$Var\left[\frac{(n-1)S^2}{\sigma^2}\right] = 2[n-1] \rightarrow \frac{(n-1)^2}{\sigma^4} V[S^2] = 2[n-1]$$

$$\rightarrow V[S^2] = \frac{2\sigma^4}{n-1}$$

Would suggest you usign new line for each diff step rather than using →
Correct derivation though
[1+2]
[3/4]

Question 2

Given,

$$n = 49\ and\ 99\%\ CI\ for\ \mu \in (30{,}50)$$

$$99\%\ CI\ for\ \mu \in (30{,}50) \rightarrow \left(30 = \bar{X} - Z_{0.005}\ \frac{\sigma}{\sqrt{49}}, 50 = \bar{X} + Z_{0.005}\ \frac{\sigma}{\sqrt{49}}\right)$$

$$\rightarrow \left(30 = \bar{X} - 2.5758\frac{\sigma}{7}, 50 = \bar{X} + 2.5758\frac{\sigma}{7}\right)$$

By solving it we get,

$$\bar{X} = 40, \sigma = 27.176$$

$$90\%\ CI\ for\ \mu \in \left(\bar{X} \pm Z_{0.05} \times \frac{\sigma}{\sqrt{49}}\right) \rightarrow \left(40 \pm 1.6449 \times \frac{27.176}{7}\right) \rightarrow (33.614, 46.385)$$

So the $90\%\ CI\ for\ \mu\ is\ (33.614, 46.385)$

Mention table book refrence for the formula as well

Correct CI though

[5/5]

Question 3

1) As we know the sampling distribution of sample variance,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$applying\ expectation\ both\ sides,$

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = [n-1] \rightarrow \frac{n-1}{\sigma^2} E[S^2] = n-1 \rightarrow E[S^2] = \sigma^2$$

2) A) Given,

$$G^2 = \frac{1}{n}[\sum x_i^2 - n\bar{x}^2]$$

$$E[G^2] = \frac{1}{n}\left[\sum E[x_i^2] - nE[\bar{x}^2]\right] = \frac{1}{n}\left[\sum(\sigma^2 + \mu^2) - n\left[\frac{\sigma^2}{n} + \mu^2\right]\right]$$

$$\frac{1}{n}[n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2] = \frac{1}{n}[n\sigma^2 - \sigma^2] = \frac{n-1}{n}\sigma^2$$

$$bias = E[G^2] - \sigma^2 (as\ given\ in\ question)$$

$$bias = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$$

B) as n increases bias decreases

[5/5]

Question 4

Given,

$$X_i \sim N\left(\frac{1}{2}\nu^2, \sigma^2\right)$$

1) By MME,

$$E[x_i] = \bar{X} \rightarrow \frac{1}{2}\nu^2 = \bar{X} \rightarrow \nu^2 = 2\bar{X}$$

$$\rightarrow \hat{v} = \sqrt{2\bar{X}}$$

2) Given,

$$n = 200, \sum X_i = 900$$

$$\hat{v} = \sqrt{2 \times \frac{900}{200}} = \pm 3$$

3) When, $\sum X_i < 0$,

$$\hat{v} \ \ it\ will\ be\ invalid$$

For $\sum X_i \geq 0$,

$$\hat{v}\ will\ be\ significant$$

Correct, but (iiii) needs a little better framed description. Make sure to check out the original answer once

[4.5/5]

Question 5

As given,

$$S^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2, where\ n = 9$$

1) $P(S^2 > \sigma^2) = P\left(\frac{S^2}{\sigma^2} > 1\right) = P\left(\frac{(n-1)S^2}{\sigma^2} > 8\right) = P(\chi_8^2 > 8) = 0.5665$

All correct equations, but the final probability calcualtion is wrong, make sure to retry it once again.

2) $P(\bar{X} > \mu \mid S^2 > \sigma^2) = P(\bar{X} - \mu > 0) = P\left(\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} > 0\right) =$

$$need\ to\ mention\ that\ since\ \bar{X} and\ S^2 are\ independent$$

Incomplete part

3) $P(\bar{X} - \mu > \sigma) = P\left(\frac{(\bar{X}-\mu)}{\frac{\sigma}{\sqrt{n}}} > 3\right) = P(Z > 3) = 1 - P(Z < 3) = 1 - 0.99865 =$

1.35×10^{-3} write as 0.00135

4) $P(\bar{X} - \mu > S) = P\left(\frac{(\bar{X}-\mu)}{\frac{S}{\sqrt{n}}} > 3\right) = P(t_8 > 3) = 0.5\% < p(x) < 1\%$

You were required to do linear interpolation as well, though not show the whole workings, but the final answer need to be stated

Mention table book refrences as well

[1.5+0.5+2+1]

[5/8]

Question 6 *unattempted*

Question 7

Here, $n = 10$ and X is $number$ of cup $correctly$ $identified$

$$X \sim Bin(10, p)$$

1) $H_0: p = 0.5$ $H_1: 0.5$

$$P(tupe\ 1\ error) = P(rejecting\ H_0\ when\ H_0\ is\ true) = P(X \geq 7\ |p = 0.5)$$

$$\rightarrow P(X \geq 7) = 1 - P(X < 7) = 1 - P(X \leq 6) = 1 - 0.8281 = 0.1719$$

2) Here,

A) $H_0: p = 0.3\ v/s\ H_1: p = 0.7$

$$P(tupe\ 2\ error) = P(accepting\ H_0\ when\ H_0\ is\ FALSE) = P(X < 7 \mid p = 0.7)$$

$$= P(X \leq 6) = 0.3508$$

B) $H_0: p = 0.2\ v/s\ H_1: p = 0.8$

$$P(tupe\ 2\ error) = P(accepting\ H_0\ when\ H_0\ is\ FALSE) = P(X < 7 \mid p = 0.8)$$

$$= P(X \leq 6) = 0.1209$$

3) $power\ of\ test = 1 - P(type\ 2\ error) = 1 - 0.3508 = 0.6482$ for p=0.7

$$power\ of\ test = 1 - P(type\ 2\ error) = 1 - 0.1209 = 0.8792$$ for p=0.8

You were not required to calcualte the power of test, but to commet over that which is missing here

Too clustered presentation, difficult to check the answer. Avoid such presentations

[3+4+0.5]

[7.5/10]

Question 8

Given,

$$X_1 \text{ and } X_2 \text{ are independent both follow } U(-\theta, \theta)$$

And,

$\hat{\theta}_1 = 3Z, Z = \max(X_1, X_2)$

1) $F(Z \le z) = P(\max(X_1, X_2) \le z) = P(X_1 \le z)P(X_2 \le z)$ *since independent*

$$F(z) \times F(z) = \frac{\left(z - (-\theta)\right)^2}{\left(\theta - (-\theta)\right)^2} = \frac{(z+\theta)^2}{4\theta^2}$$

$$f(z) = \frac{dF[z]}{dz} = \frac{1}{4\theta^2} 2 \times (z + \theta) = \frac{z+\theta}{2\theta^2}$$

2) $E[Z] = \int_{-\theta}^{\theta} z \times f_z(z)\, dz = \frac{1}{2\theta^2}\int_{-\theta}^{\theta} z^2 + z\theta \; dz = \frac{1}{2\theta^2}\left[\frac{Z^3}{3} + \frac{z^2\theta}{2}\right]_{-\theta}^{\theta}$

$$\rightarrow \frac{1}{2\theta^2}\left[\frac{\theta^3}{3} + \frac{\theta^3}{2} - \left[-\frac{\theta^3}{3} + \frac{\theta^3}{2}\right]\right] = \frac{1}{2\theta^2} \times \frac{\theta^3}{3} = \frac{\theta}{3}$$

3) $E\left[\hat{\theta}_1\right] = E[3Z] = 3 \times \frac{\theta}{3} = \theta$

$$bais = E\left[\hat{\theta}\right] - \hat{\theta} = \theta - \hat{\theta} = 0$$

$$MSE\left(\hat{\theta}_1\right) = Var\left(\hat{\theta}_1\right) = 9 \times Var(Z)$$

$$E[Z^2] = \int_{-\theta}^{\theta} z^2 \times f_z(z)\, dz = \frac{1}{2\theta^2}\int_{-\theta}^{\theta} z^3 + z^2\theta \; dz = \frac{1}{2\theta^2}\left[\frac{Z^4}{4} + \frac{z^3\theta}{3}\right]_{-\theta}^{\theta}$$

$$E[Z^2] = \frac{\theta^2}{3}$$

$$Var[Z] = E[z^2] - (E[z])^2 = \frac{\theta^2}{3} - \frac{\theta^2}{9} = \frac{2\theta^2}{9}$$

So,

$$MSE\left(\hat{\theta}_1\right) = Var\left(\hat{\theta}_1\right) = 9 \times Var(Z) = 2\theta^2$$

Subpart a and b not metioned in iii question, no marks to be allotted. Correct solution though

4) Now, $\hat{\theta}_2 = 2Z$

$$E\left[\hat{\theta}_2\right] = 2E[Z] = 2 \times \frac{\theta}{3} = \frac{2\theta}{3}$$

$$bais = E[\hat{\theta}_2] - \hat{\theta} = \frac{2}{3}\hat{\theta} - \hat{\theta} = -\frac{\hat{\theta}}{3}$$

$$MSE[\hat{\theta}_2] = Var(\theta_2) + bais^2 = 4 \times V[Z] + \frac{\hat{\theta}^2}{9} = \frac{8\theta^2}{9} + \frac{\theta^2}{9} = \frac{9\theta^2}{9} = \theta^2$$

Same issue as in (iii), no subpart mentioned, thus no marks allotted .Correct solution though

5) $\widehat{\theta_1}$ *is an unbaised parameter and have MSE OF* $2\theta^2$ *wherease the* $\hat{\theta}_2$ *is not unbaised but have MSE of* θ^2 *which is lower than* θ_1 .

correct interpretation

[4+1+0+0+2]

[7/16]

question 9

let,

$survey\ 1,2,3 = 1,2,3\ respectively$

$$propotion\ of\ smoker\ in\ survey\ 1 = \hat{p}_1 = \frac{36}{124} = 0.2903$$

$$propotion\ of\ smoker\ in\ survey\ 2 = \hat{p}_2 = \frac{25}{136} = 0.1838$$

1) $H_0: p_1 = p_2\ \ v/s\ \ H_1: p_1 \neq p_2$

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} = 0.2346$$

$TSV\ under\ H_0,$

$$TSV = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.1065}{\sqrt{\frac{0.1795}{124} + \frac{0.1795}{136}}} = 2.2024$$

$$CV = Z_{0.025} = 1.96$$

$we\ are\ carrying\ out\ a\ 2\ tail\ test\ at\ 5\%\ significant\ level\ of\ propotion\ equality,$

$here\ we\ can\ see\ that\ |TSV| > CV, so\ there\ is\ sufficent\ evidence\ to\ reject\ H_0, so\ there\ is$

$significant\ diffrence\ between\ the\ smoker\ propotion\ in\ both\ surveys$

Correct

Would suggest you to use normal text (not eq editor) to write the text portion

2) Contingency table

Here,

Survey	smoker	Non smoker	total
1	36	88	124
2	25	111	136
3	26	115	141
total	87	314	401

$$H_0: propotion\ is\ diffent\ in\ all\ three\ surveys$$

$$H_1: proportion\ is\ not\ diffrent\ in\ all\ three\ surveys$$

E_i ,

26.90	97.09
29.506	106.49
30.59	110.40

$$TSV = C_i = \frac{\sum(o_i - E_i)^2}{E_i}$$

By solving,

$$TSV = 5.689$$

$$CV = \chi^2_{(3-1)(2-1),0.05} = \chi^2_{2,0.05} = 5.991$$

We are carrying out one tail contingency table statistical test at 5% significant level, since

$TSV < CV$, so we don't have sufficient evidence to reject
$H_O\ so\ there\ is\ significant\ diffrence\ btw\ all\ three\ smoker\ propotion$

Need to specifically mention out about the dof of chi-square test. 1 mark deducted for the same.

3) A) $\hat{p}_3 = \frac{26}{141} = 0.1843$

B)

$both\ the\ test\ prooves\ that\ there\ is\ diffrence\ btw\ the\ propotion\ of\ smoker\ in\ all\ three$
Given survey. Not a satisfactory explanation, read out the original one

[7+6+1]
[14/16]

TOTAL:51/77

Would have received more marks if numbering would have been correct
Good performance overall, few areas of improvisation as marked specifically for each question, make sure to work on that.
Also one question left unattempted, If it is because of time constraint, make sure to speed up the pace of attempting the paper.