Q2 1). If calculate the mean with class number equals to number of class a student will take,

the mean $= \frac{100 + 100 + 10}{3} = 70$

If calculate the mean with total class number,

the mean $= \frac{100 + 100 + 10 \times 10}{2 + 10} = 25$

$\therefore$ So both sides are right

2) i) Let the sample proportion be p.

$\alpha = 0.1$

$z_{\frac{\alpha}{2}} = 1.645$

$p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} = [0.8265, 0.9095]$

$p = 0.868$

Sampling error $= E = z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} = 0.0415$

sample size used $= \frac{(z_{\frac{\alpha}{2}})^2 p(1-p)}{E^2}$

$= \frac{1.645^2 (0.868)(1-0.868)}{0.0415^2}$

$= 180.0235 \approx 181$

ii). Sample size needed $= \frac{1.645^2 (0.5)(1-0.5)}{0.06^2}$

$\because$ no past data

$\therefore \pi \approx 0.5$

$= 187.9184$

≈ 188

iii). $H_0: \pi \geq 0.9$ ∵ As $n = 200 > 30$ ∴ $p \sim N$

$H_1: \pi < 0.9$.

$np = 200 - 21 = 179 > 5$

$n(1-p) = 21 > 5$

at $\alpha = 0.1$

$Z_c = \pm 1.285$

$p = \frac{179}{200} = 0.895$.

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0.895 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{200}}} = -0.2357$$

∵ $-0.2357 > -1.285$

∴ We do not reject H_0

∴ There is insufficient evidence that the awareness rate in HK is lower than the target awareness rate 0.9.

iv) ∵ $H_0: \pi \geq 0.9$, $H_1: \pi < 0.9$.

Type I error α: It is the probability that H_0 is rejected given that it is true.

Type II error β: It is the probability that H_0 is not rejected given that it is false.

LQ2) ai). It is the best estimation to an unknown population value.

ii). Assumption: The population is normal, no. of student ≥ 30

$\bar{x} = 74.5$, $s = 6.2335$, $n = 8$, d.f. $= 7$

$$90\% \text{ C.I.} = \bar{x} \pm Z_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 74.5 \pm 1.645\left(\frac{6.2335}{\sqrt{8}}\right)$$

$$= [70.8746, 78.1254]$$

2a iii). Assumption: population is normal, No. of student ≥ 30

$H_0: \mu = 78.5$ $\quad \bar{x} = 74.5$ $\quad n=8$

$H_1: \mu \neq 78.5$ $\quad S = 6.2335$ $\quad d.f. = 7$

$\alpha = 0.1$

$z_c = \pm 1.645$

$z = \dfrac{\bar{x}-\mu}{s/\sqrt{n}}$

$= \dfrac{74.5-78.5}{6.2335/\sqrt{8}}$

$= -1.8150$

$\because -1.8150 < -1.1645$

$\therefore$ Reject H_0

$\therefore$ There is sufficient evidence that the population mean is not 78.5

iv). p-value $= 2P(z \leq -1.8150)$
$= 0.0695$

v). Yes.
Because with 90% C.I. $= [70.8746, 78.1254]$,
78.5 is not in that range.

bi) $S_{uu} = 280\,000$

$S_{uu} = 1150$

$S_{uu} = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} = 16500.$

$r = \dfrac{S_{uu}}{\sqrt{S_{uu}S_{uu}}} = \dfrac{16500}{\sqrt{280000 \times 1150}} = 0.9195$

ii). Let $Y = A + Bx$

$B = \dfrac{S_{uu}}{S_{uu}} = \dfrac{16500}{280000} = 0.0589$

$A = \bar{y} - B\bar{x}$

$= 60 - 0.0589(400)$

$= 36.4286$

$Y = 36.4286 + 0.0589x$

with slope $= 0.0589$, intercept $= 36.4286$

IV) 1) weight = 36.4286 + 0.0589 (650)

= 74.7321 cm

2) weight = 36.4286 + 0.0589 (1000)

= 95.3571 cm

V). (1), the 650 kg is more justifiable

∵ The weight 1000 kg is outside the relevant range for the independent variable. So, it is not appropriate to use the model to predict the length of tail

III) var = 0.000127

$\sigma = 0.0113$

$$SSE = \frac{S_{ww}S_{uu} - S_{wu}^2}{S_{ww}}$$

$$= \frac{280000(1150) - 16500^2}{280000}$$

$$= 177.6786$$

$$S_{b_1} = \sqrt{\frac{SSE/(n-2)}{S_{ww}}}$$

$$= \sqrt{\frac{177.6786/(7-2)}{280000}}$$

$$= 0.0113$$

$$90\%\ C.I. = B \pm t_{0.05,5} S_{b_1} = 0.0589 \pm 1.9432(0.0113)$$

$$= [0.0370, 0.0808]$$