S3 June 2018 (MA)

Q1 a) $r = \dfrac{S_{ca}}{\sqrt{S_{cc} \times S_{aa}}} = \dfrac{47.7625}{\sqrt{34787.5 \times 0.2172875}} = \boxed{0.549}$

b) $H_0: \rho = 0$ critical value: ± 0.6215 (5%, 1-tail)

$H_1: \rho > 0$

$0.549 < 0.6215$

$\therefore$ Result is insignificant.
Accept H_0.
Evidence suggests no +ve correlation exists between concentration of radioactive element and amount of dissolved solids in groundwater.

c)

Sample	c	a	R_c	R_a	d	d^2
A	625	1.28	2	4	2	4
B	700	1.30	6	5	1	1
C	650	1.00	4	1	3	9
D	645	1.20	3	3	0	0
E	720	1.55	7	8	1	1
F	600	1.15	1	2	1	1
G	825	1.40	8	6	2	4
H	665	1.45	5	7	2	4
						24

$\sum d^2 = 24 \therefore r_s = 1 - \frac{6(24)}{8(63)}$

$= \boxed{0.714}$

d) $H_0 : \rho = 0$ critical value : 0.6429
$H_1 : \rho > 0$ (5%, 1-tail)

$0.714 > 0.6429$

$\therefore$ Result is significant.
Reject H_0.
Evidence suggests a +ve correlation exists between c and a.

e) Data likely not jointly normally distributed, which is required to use PMCC.

Q2a) - List all ticket numbers for standard/premium
- Use random numbers to select a sample for both standard and premium ticket holders.
- These sample sizes must be proportionate to the total population (ie all ticket holders at the concert).

b) $H_0: \mu_p - \mu_s = 6$ where $\left[\begin{array}{l} p = \text{premium} \\ s = \text{standard} \end{array}\right]$

$H_1: \mu_p - \mu_s > 6$

$$\text{Test Statistic} = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

$$= \frac{23 - 15 - (6)}{\sqrt{\frac{10^2}{80} + \frac{8^2}{55}}}$$

$$= 1.189\ldots$$

critical value: ± 1.6449
(5%, 1-tail)

So ... $1.189 < 1.6449$

$\therefore$ Result is insignificant.

Accept H_0.

Evidence suggests manager's claim is incorrect.

c) No - sample sizes are large so C.L.T applies.

(Q3a) $\sum x = 6.2$, $\sum x^2 = 9.78$, $n = 4$

$$\text{mean} = \frac{\sum x}{n} = \frac{6.2}{4} = \boxed{1.55}$$

$$\hat{\text{variance}} = s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{(n)}\right)$$

$$= \frac{1}{3}\left(9.78 - \frac{(6.2)^2}{4}\right)$$

$$= \boxed{0.057}$$

b) $T \sim N(\mu, 0.5^2)$ where T = repair time

$\bar{T}$ = estimate of population mean...

$\bar{T} \sim N\left(\mu, \frac{0.5^2}{n}\right)$ By C.L.T

P(required) $\Rightarrow P(|\bar{T} - \mu| < 0.1) \geq 0.99$

$$P(|\bar{T} - \mu| < 0.1) = P(-0.1 < \bar{T} - \mu < 0.1) \geq 0.99$$

$$P(-0.1 + \mu < \bar{T} < 0.1 + \mu) \geq 0.99$$

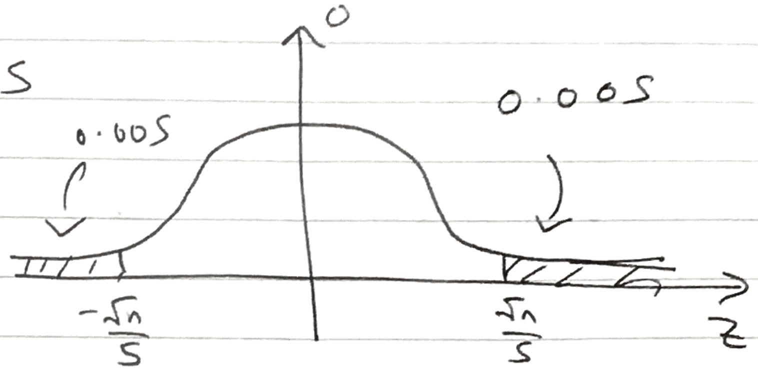
Z

$$P\left(\frac{-0.1 + \mu - \mu}{\frac{0.5}{\sqrt{n}}} < Z < \frac{0.1 + \mu - \mu}{\frac{0.5}{\sqrt{n}}}\right) \geq 0.99$$

$$P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) \geq 0.99$$

Because Area of unshaded region ≥ 0.99!

so $P\left(Z > \frac{\sqrt{n}}{5}\right) \leq 0.005$



but $P(Z > 2.5758) = 0.005$

this means that $\frac{\sqrt{n}}{5} \geq 2.5758$

$\sqrt{n} \geq 5 \times 2.5758$

$n \geq (5 \times 2.5758)^2$

$n \geq 165.87\ldots$

so $\boxed{n_{min} = 166}$

(Q4a) width $= 2(2.5758)\frac{\sigma}{\sqrt{n}} = 2(2.5758)\frac{(\sigma)}{\sqrt{120}}$

'critical' value

$= \boxed{0.470\sigma}$

b) $H_0: \mu = 6$
$H_1: \mu \neq 6$ } 6 is in the interval!

so Result is insignificant
Accept $H_0 \rightarrow \mu = 6$

sig level: 10%

c) width of given interval $= 6.25 - 5.14$
$= 1.11$

width of given interval $= 2(1.6449)\left(\frac{\sigma}{\sqrt{100}}\right)$

$[P(Z > 1.6449) = 0.05]$

so $\frac{2(1.6449)\sigma}{10} = 1.11$

$\sigma = \frac{1.11 \times 10}{2 \times 1.6449} = \boxed{3.37}$

Q5a) $C \sim N(1000, 250^2)$

$L \sim N(2800, 650^2)$

$P(\text{required}) = P(L > 3C) = P(L - 3C > 0)$

let $A = L - 3C$, $E(A) = 2800 - 3(1000)$
$= -200$ //

$Var(A) = Var(L) + 9Var(C)$
$= 650^2 + 9(250^2)$
$= 985000$ //

$\therefore A \sim N(-200, 985000)$

$P(A > 0) = P\left(Z > \frac{200}{\sqrt{985000}}\right) = P(Z > 0.20)$

$= 1 - P(Z < 0.20) = \boxed{0.4207}$

b) let $B = 8C + 3L$,

$E(B) = 8(1000) + 3(2800) = \underline{\underline{16400}}$

$Var(B) = 8^2 Var(C) + 3^2 Var(L)$

$= 64(250^2) + 3^2(650^2) = \underline{\underline{1767500}}$

$P(\text{required}) = P(B > 20000)$

$= P\left(Z > \dfrac{20000 - 16400}{\sqrt{1767500}}\right)$

$= P(Z > 2.71) = 1 - P(Z < 2.7)$

$\simeq \boxed{0.0035}$

c) - selection of cars/lorries is <u>random</u>

<u>OR</u> - weights of cars/lorries are independent.

$X \sim B[4, 0.5]$

(Q6a) H_0: $B[4, 0.5]$ is a suitable model.

H_1: $B[4, 0.5]$ is not a suitable model.

No.	0	1	2	3	4
O_i	12	45	36	39	18
E_i	$150 \times P(X=0)$	$150 \times P(X=1)$	$150 \times P(X=2)$	$150 \times P(X=3)$	$150 \times P(X=4)$

$E_i = 150 \times P(X = i)$ //

because expeiment is repeated 150 times.

No.	0	1	2	3	4
O_i	12	45	36	39	18
E_i	9.375	37.5	56.25	37.5	9.375
$\frac{(O-E)^2}{E}$	0.735	1.50	7.29	0.06	7.935

$$X^2 = \sum \frac{(O-E)^2}{E} = 17.5 \text{ //}$$

$\gamma = 5 - 1 = 4$ $\therefore$ critical value $= \chi^2_4(1\%) = 13.277$ //

$17.5 > 13.277$

$\therefore$ Result is significant

Reject H_0

$B[4, 0.5]$ isn't a suitable model!

b) $\hat{p} = \dfrac{\text{total dices showing even no.}}{\text{total dice rolls}}$

$= \dfrac{0(12)+1(45)+2(36)+3(39)+4(18)}{4 \times 150}$

$= \boxed{0.51}$

c) $d = 150P(X=2) = 150 \times \left[\binom{4}{2}(0.51)^2(1-0.5))^2\right]$

$= \boxed{56.2}$

$e = 150 - (\Sigma' f_i) = \boxed{10.2}$

d) H_0: $B[4, 0.51]$ is a suitable model

H_1: $B[4, 0.51]$ is not a suitable model.

e) $\nu = 5-1-1 = 3 \rightarrow \chi^2_3(1\%) = 11.345$

(Subtract an extra 1 as p was calculated.)

$16.9 > 11.345$

$\therefore$ Result is significant.

Reject H_0

$B[4, 0.51]$ is not a suitable model!