S3 June 2015 IAL (MA)

Q1a) $\frac{720}{40} = \boxed{18}$

b) $\boxed{18}$

c) 0 - first name is chosen randomly from first 18 members

d) ADV : - Simple and easy to use
- Suitable for large samples
- Gives a good spread of data

any 1

DISADV : - alphabetical list is not random so bias is introduced.
- some combinations of names won't be possible

either

Q2a)

Dancer	Judge 1 Rank	Judge 2 Rank	d	d^2
A	6	8	2	4
B	3	4	1	1
C	4	5	1	1
L	9	7	2	4
N	2	3	1	1
R	8	9	1	1
S	1	1	0	0
T	5	2	3	9
Y	7	6	1	1
				22

$$\sum d^2 = 22$$

$$\therefore r_s = 1 - \frac{6(22)}{9(80)} = \boxed{0.817}$$

b) $H_0: \rho = 0$ critical value: ± 0.7833

$H_1: \rho > 0$

agreement indicated by +ve correlation

$0.817 > 0.7833$

$\therefore$ Result is significant – Reject H_0.

Evidence suggests judges are generally in agreement.

Q3a) mean = $\frac{\text{total accidents}}{\text{total days}}$

$= \frac{0(47) + 1(57) + 2(46) + 3(35) + 4(9) + 5(6)}{200}$

$= \boxed{1.6}$

b) $r = 200 \times P(X=2) = 200 \left[\frac{e^{-1.6}(1.6^2)}{2!} \right] = \boxed{51.69}$

for 200 days.

$s = 200 - (\sum E_i) = \boxed{4.72}$

c) H_0: Poisson ($\lambda = 1.6$) is a suitable model for these data

H_1: Poisson ($\lambda = 1.6$) is not a suitable model for these data.

Remember, expected frequencies must be greater than 5 for the test statistic X^2 to be approximated well by the chi-squared distribution(s).

So pool final two groups:

No. accidents	0	1	2	3	≥ 4
E_i	40.38	64.61	51.69	27.57	15.75
O_i	47	57	46	35	15
$\frac{(O-E)^2}{E}$	1.0853	0.8963	0.6264	2.0024	0.0357

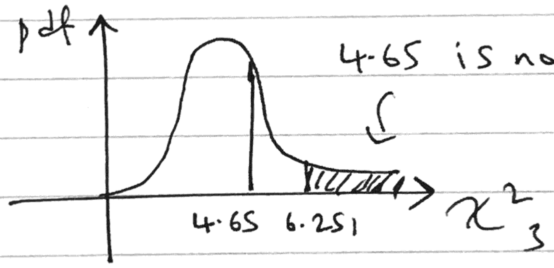
$$X^2 = \sum_i \frac{(O-E)^2}{E} = 4.65$$

$$\nu = 5 - 1 - 1 = 3$$

subtract an additional 1 as the parameter λ was calculated

$\therefore$ critical value $= \chi^2_3(10\%) = 6.251$

$$4.65 < 6.251$$



4.65 is not in the C.R.

$\therefore$ Result is insignificant.
Accept H_0.
Evidence suggests that Poisson is a suitable model - supervisor's belief is correct.

Q4a) let S = weight of a sack of potatoes.

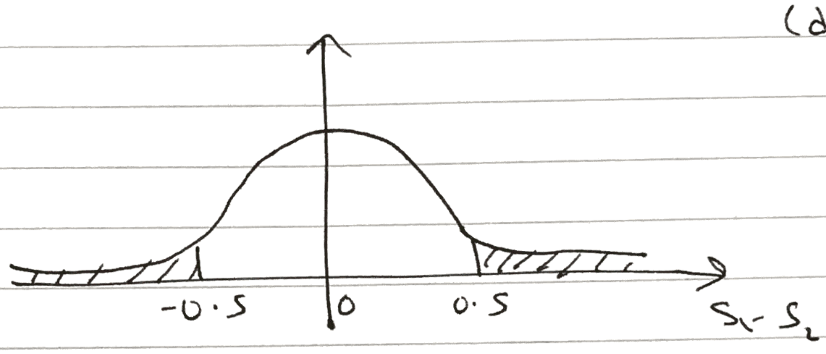
$$S \sim N(25.6, 0.24^2)$$

$$P(\text{required}) = P(|S_1 - S_2| > 0.5)$$

$$(S_1 - S_2) \sim N(0, 2(0.24^2))$$

$$P(|S_1 - S_2| > 0.5) = 2P(S_1 - S_2 > 0.5)$$

(due to symmetry)



$$= 2\left[P\left(Z > \frac{0.5 - 0}{\sqrt{2(0.24^2)}}\right)\right] = 2P(Z > 1.47)$$

$$= 2\left[1 - P(Z < 1.47)\right] = 2(1 - 0.9292)$$

$$= \boxed{0.1416}$$

b) let P = pallet (empty), $P \sim N(20, 0.32^2)$

let Full pallet $= (S_1 + \ldots + S_{30} + P) = F$

$$E(F) = 30E(S) + E(P) = 30(25.6) + 20 = \underline{\underline{788}}$$

$$Var(F) = 30\,Var(S) + Var(P) = 30(0.24^2) + 0.32^2$$

$$= \underline{\underline{1.8304}}$$

So $F \sim N(788, 1.8304)$

$$P(\text{required}) = P(F > 785)$$

$$= P\left(Z > \frac{785 - 788}{\sqrt{1.8304}}\right)$$

$$= P(Z > -2.22)$$

$$= P(Z < 2.22) = \boxed{0.9868}$$

Q5) H_0: There is no association between Grade and Gender.

H_1: There is an association between Grade and Gender.

working out actual observed values:

Male / Distinction = $200 \times 0.185 = 37$
Male / Merit = $200 \times 0.635 = 127$
Male / Unsatisfactory = $200 \times 0.180 = 36$

Female / Distinction = $160 \times 0.275 = 44$
Female / Merit = $160 \times 0.600 = 96$
Female / Unsatisfactory = $160 \times 0.125 = 20$

OBSERVED:

	Male	Female
Distinction	37	44
Merit	127	96
Unsatisfactory	36	20

$$\text{expected no.} = \frac{\text{Row Total} \times \text{Column Total}}{\text{Grand Total}}$$

EXPECTED :

	Male	Female	
Distinction	45	36	81
Merit	123.89	99.11	223
Unsatisfactory	31.11	24.89	56
	200	160	$\boxed{360}$

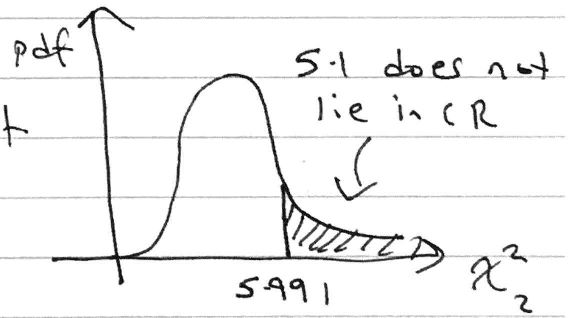
O_i	E_i	$\frac{(O-E)^2}{E}$
37	45	1.42
44	36	1.78
127	123.89	0.08
96	99.11	0.10
36	31.11	0.77
20	24.89	0.96
		5.11

$$X^2 = \sum \frac{(O-E)^2}{E} = 5.1$$

$$\nu = (\text{rows} - 1)(\text{columns} - 1) = (3-1)(2-1) = 2$$

$\therefore$ critical value $= \chi^2_2(5\%) = 5.991$

$5.1 < 5.991$



$\therefore$ Result is insignificant
Accept H_0
Evidence suggests Head of Department's belief is correct.
(No association)

Q6a) mean $= \frac{\sum x}{n} = \frac{1570}{50} = \boxed{31.4} = \bar{x}$

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

$= \frac{1}{49}\left(49467.58 - \frac{(1570)^2}{50}\right)$

$= \boxed{3.46}$

b) $H_0: \mu_X = \mu_Y$ where X refers to morning sample and Y refers to late afternoon.

$H_1: \mu_X > \mu_Y$

critical value: ± 1.6449 (5%, 1-tail)

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

$= \frac{31.4 - 30.9}{\sqrt{\frac{3.46}{50} + \frac{3.03}{50}}}$ s^2 from (a)

$= 1.39$

$1.39 < 1.6449$

$\therefore$ Result is insignificant.

Accept H_0 - no difference in mean times

c) Allows us to assume $\bar{X}$ and $\bar{Y}$ (sample means) are normally distributed as n is large.

d) sample variance = population variance
($s^2 = \sigma^2$)
(By definition the test statistic uses σ^2).

Q7a) (Discrete Uniform Distribution)

s	1	2	3	4	5	6
$P(S=s)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(S) = \frac{a+b}{2} = \frac{1+6}{2} = \boxed{3.5}$

$Var(S) = \frac{n^2-1}{12} = \frac{6^2-1}{12} = \boxed{\frac{35}{12}}$

alt: use $E(S^2) - [E(S)]^2 = Var(S)$ //

b) $P(\bar{S} < 3) = P(\text{required})$

$\bar{S} \sim N\left(3.5, \frac{35}{12(40)}\right)$, By C.L.T //

$\Rightarrow P(\bar{S} < 3) = P\left(Z < \frac{3-3.5}{\sqrt{\frac{35}{12(40)}}}\right)$

$= P(Z < -1.85) = 1 - P(Z < 1.85)$

$= \boxed{0.0322}$

Q8a) $\bar{x} = \dfrac{29.74 + 31.86}{2} = 30.8$

standard error $= \dfrac{\sigma}{\sqrt{n}}$

$\bar{x} + (1.96)\dfrac{\sigma}{\sqrt{n}} = 31.86$

$1.96\dfrac{\sigma}{\sqrt{n}} = 31.86 - \bar{x}$

$\therefore \dfrac{\sigma}{\sqrt{n}} = \dfrac{31.86 - 30.8}{1.96} = \boxed{0.54}$

b) 90% C.I: $\left[\bar{x} \pm 1.6449\dfrac{\sigma}{\sqrt{n}}\right]$

$\left[30.8 \pm 1.6449(0.54)\right]$

$\left[29.91,\ 31.69\right]$

c) $X \sim B[4, 0.90]$ where X = no. of confidence intervals containing μ. (out of 4).

$P(X \geq 3) = P(X=3) + P(X=4)$

$= \binom{4}{3}(0.9^3)(0.1) + 0.9^4 = \boxed{0.948}$