

S3 June 2014 (MA)

- Q1a) This is a sample where every other possible sample of size n has an equal chance of being selected.
- b) When it's not possible to obtain a sampling frame.
- c) i) A list of all the students.
 ii) Create a numbered list of all students (1-75). Then use random numbers to select 8 students. Pick only numbers in the range 1-75 and the numbers picked will correspond to the students in the list.

Q2a) Yes - no population parameters included, only known data

- ii) No - contains μ , a population parameter.
- iii) No - contains μ and σ , both population parameters.

$$b) E\left(\frac{3X_1 - X_{20}}{2}\right) = \frac{1}{2} [E(3X_1 - X_{20})]$$

$$= \frac{1}{2} [3E(X_1) - E(X_{20})]$$

$$= \frac{1}{2} [3\mu - \mu] = \mu \quad \text{(unbiased estimator of } \mu \text{)}$$

$$\begin{aligned} \text{Var} \left(\frac{3X_1 - X_{20}}{2} \right) &= \frac{1}{4} \text{Var} (3X_1 - X_{20}) \\ &= \frac{9}{4} \text{Var} (X_1) + \frac{1}{4} \text{Var} (X_{20}) \\ &= \frac{9}{4} \sigma^2 + \frac{1}{4} \sigma^2 = \boxed{\frac{5\sigma^2}{2}} \end{aligned}$$

Q3) H_0 : There is no association between gender and happiness.

H_1 : There is an association between gender and happiness.

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

EXPECTED :	NH	FH	VH
F	13.51	41.77	30.71
M	8.49	26.23	19.29

O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
9	13.51	1.508
43	41.77	0.0361
34	30.71	0.351
13	8.49	2.402
25	26.23	0.0575
16	19.29	0.560

$$\chi^2 = \sum \frac{(O - E)^2}{E} = 4.91$$

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

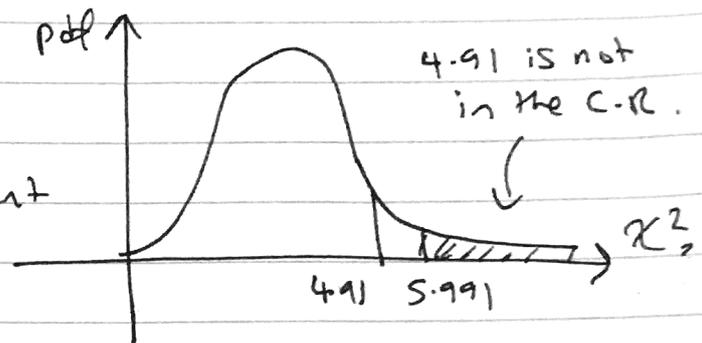
$$\therefore \text{critical value} = \chi^2_2 (5\%) = 5.991$$

$$\therefore \text{Result} \quad 4.91 < 5.991$$

is insignificant

Accept H_0 .

Evidence suggests
gender is independent
of happiness.



(04)

$$A = B + 4C - 3D$$

$$E(A) = E(B) + 4E(C) - 3E(D)$$

$$= 6 + 4(7) - 3(4) = 22$$

$$\text{Var}(A) = \text{Var}(B) + 4^2 \text{Var}(C) + 3^2 \text{Var}(D)$$

$$= 2^2 + 4^2(3^2) + 3^2(1.5^2)$$

$$= 168.25$$

$$\therefore A \sim N(22, 168.25)$$

$$P(A < 45) = P\left(Z < \frac{45 - 22}{\sqrt{168.25}}\right)$$

$$= P(Z < 1.77) = \boxed{0.9616}$$

- (Q5a) - fixed number of seeds in a row.
 - only two outcomes \rightarrow either a seed germinates or it does not.

$$b) p = \frac{\text{no. of germinated seeds}}{\text{total no. of seeds}}$$

$$= \frac{0(2) + 1(6) + 2(11) + 3(19) + 4(25) + 5(32) + 6(16) + 7(9)}{120 \times 7}$$

$$= \frac{504}{840} = \boxed{0.6}$$

$$c) s = 120 \times P(X=2) \quad \text{where } X \sim B[7, 0.6]$$

$X = \text{no. of seeds germinating in one row.}$

$$= 120 \left[\binom{7}{2} (0.6)^2 (0.4)^5 \right]$$

$$= \boxed{9.29}$$

$$t = 120 - (\sum E_i) = \boxed{34.84}$$

H_0 : Binomial is a suitable model
 H_1 : Binomial is not a suitable model.

d) Remember that expected frequencies must all be greater than 5 for the test statistic χ^2 to be approximated well by the chi-squared distribution (χ^2).

So pool first 3 cells and last 2...

No. of ...	0-2	3	4	5	6-7
E_i	11.55	23.22	34.84	31.35	19.04
O_i	19	19	25	32	25
$\frac{(O-E)^2}{E}$	4.865	0.767	2.779	0.013	1.866

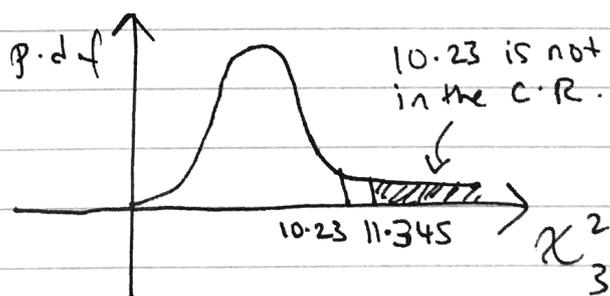
$$\chi^2 = \sum \frac{(O-E)^2}{E} = 10.23 //$$

$$\gamma = 5 - 1 - 1 = 3$$

[subtract an extra '1' as we calculated p]

$$\therefore \text{critical value} = \chi^2_3(1\%) = 11.345$$

$$10.23 < 11.345$$



\therefore Result is insignificant.

Accept H_0 .

Binomial is a suitable model.

Q6a) if \bar{X} is an unbiased estimator of μ
then $E(\bar{X}) = \mu$.

$$E(\bar{X}) = E\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right)$$

$$= \frac{1}{n}[E(X_1) + E(X_2) + \dots + E(X_n)]$$

$X_1, X_2, X_3, \dots, X_n$ all have an expected value of μ .

$$\therefore \frac{1}{n}[\mu + \mu + \dots + \mu] = E(\bar{X}) = \frac{n \times \mu}{n}$$

$$= \mu$$

b) $\frac{\sum x}{n} = \text{mean} = \frac{1001}{5} = \boxed{200.2}$

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

$$= \frac{1}{4} \left(200.469 - \frac{(1001)^2}{5} \right)$$

$$= \boxed{17.2}$$

c) \bar{X} = estimated population mean.

what we want is ...

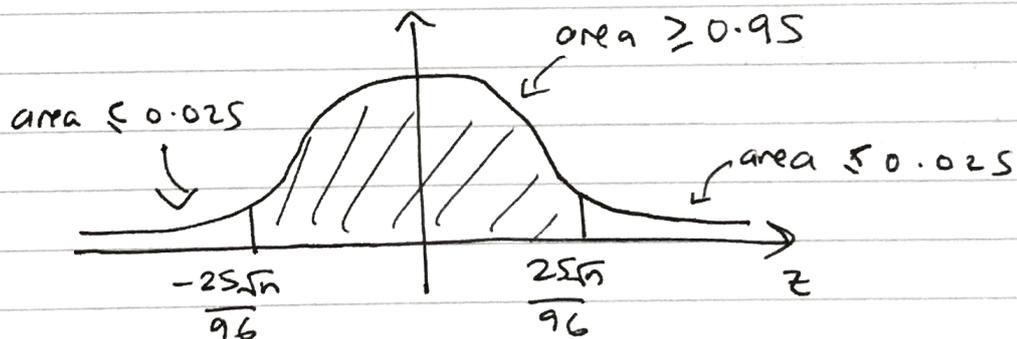
$$P(-1.25 < \bar{X} - \mu < 1.25) \geq 0.95$$

$$P(-1.25 + \mu < \bar{X} < 1.25 + \mu) \geq 0.95$$

Now by C.L.T, $\bar{X} \sim N\left(\mu, \frac{4.8^2}{n}\right)$

$$P\left(\frac{-1.25 + \mu - \mu}{\frac{4.8}{\sqrt{n}}} < Z < \frac{1.25 + \mu - \mu}{\frac{4.8}{\sqrt{n}}}\right) \geq 0.95$$

$$P\left(-\frac{25\sqrt{n}}{96} < Z < \frac{25\sqrt{n}}{96}\right) \geq 0.95$$



the 'middle' area ≥ 0.95 \therefore the area to the right (or left) ≤ 0.025

so we can say: $P\left(Z > \frac{25\sqrt{n}}{96}\right) \leq 0.025$

the z-value corresponding to $(p=0.025) = 1.96$

$\therefore \frac{25\sqrt{n}}{96} \geq 1.96$ (as z-value increases, $P(Z > z)$ decreases)

$$\text{so } \sqrt{n} \geq \frac{1.96(96)}{25}$$

$$n \geq 56.6 \dots$$

$$\therefore \boxed{n_{\min} = 57} \quad (\text{next integer})$$

$$\text{Q7a)} \quad P(X < \mu - 30) = 0.0005$$

$$P\left(z < \frac{\mu - 30 - \mu}{\sigma}\right) = 0.0005$$

$$P\left(z > \frac{30}{\sigma}\right) = 0.0005.$$

$$\therefore \text{from tables, } 3.2905 = \frac{30}{\sigma}$$

$$\sigma = \frac{30}{3.2905} = \boxed{9.117g}$$

$$\text{b)} \quad H_0: \mu = 1000$$

$$H_1: \mu < 1000$$

$$\bar{x} = \frac{\sum x}{n} = 999.54$$

critical value: ± 2.3263
(1%, 1-tail)

$$\text{Test Statistic} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{999.54 - 1000}{\frac{9.117}{\sqrt{10}}}$$

$$= -0.160 //$$

$$-0.160 > -2.3263$$

\therefore Result is insignificant

Accept H_0 . Mean is not less than 1kg.

$$d) H_0: \rho = 0$$

$$H_1: \rho > 0$$

↑
↓

ability to correctly
order men would
mean a positive correlation.

critical value: 0.600
(5%, 1-tail)

$$0.417 < 0.600$$

∴ Result is ~~sig~~ insignificant.

Accept H_0 .

Peter does not have
the stated ability.