Probability & Statistics
Mid Solution (Spring 2021)

Sol 1 (a) Consider three slots $\underset{1}{\boxed{}}\underset{2}{\boxed{}}\underset{3}{\boxed{}}$ which corresponds to three positions.
There are total $40 \times 39 \times 38 = 59280$ ways in which we can fill these three slots from boys and girls. Only in $35 \times 34 \times 33 = 39270$ ways we can fill these three slots from boys.
Thus the probability that boys will receive top three positions is

$$\frac{39270}{59280} = 0.66$$

(b) Let P_y is the event that a student know Python and M is the event that a student know MATLAB. Then it is given that $P(P_y) = 90/100$, $P(M) = \frac{20}{100}$ and $P(P_y \cap M) = 5/100$.

Required $P(M \mid P_y)$

We know that $P(M \mid P_y) = \dfrac{P(M \cap P_y)}{P(P_y)} = \dfrac{5/100}{90/100} = \dfrac{5}{90}$
$= \frac{1}{18} \approx 0.06$

Sol 2 (a) The sample correlation coefficient 'r' measure the linear relation b/w the x and y values of a paired data set. If $|r|=1$ then there will be a perfect ~~relati~~ linear relation b/w x and y values — this mean that a straight line can pass through our data points (x_i, y_i), $i=1,2,\ldots,n$. If r is positive then smaller y values tend to go with smaller x values and larger y values with larger x values. If r is negative then larger y values tend to go with smaller x values and smaller y values with larger x values. Furthermore $-1 \leq r \leq 1$.

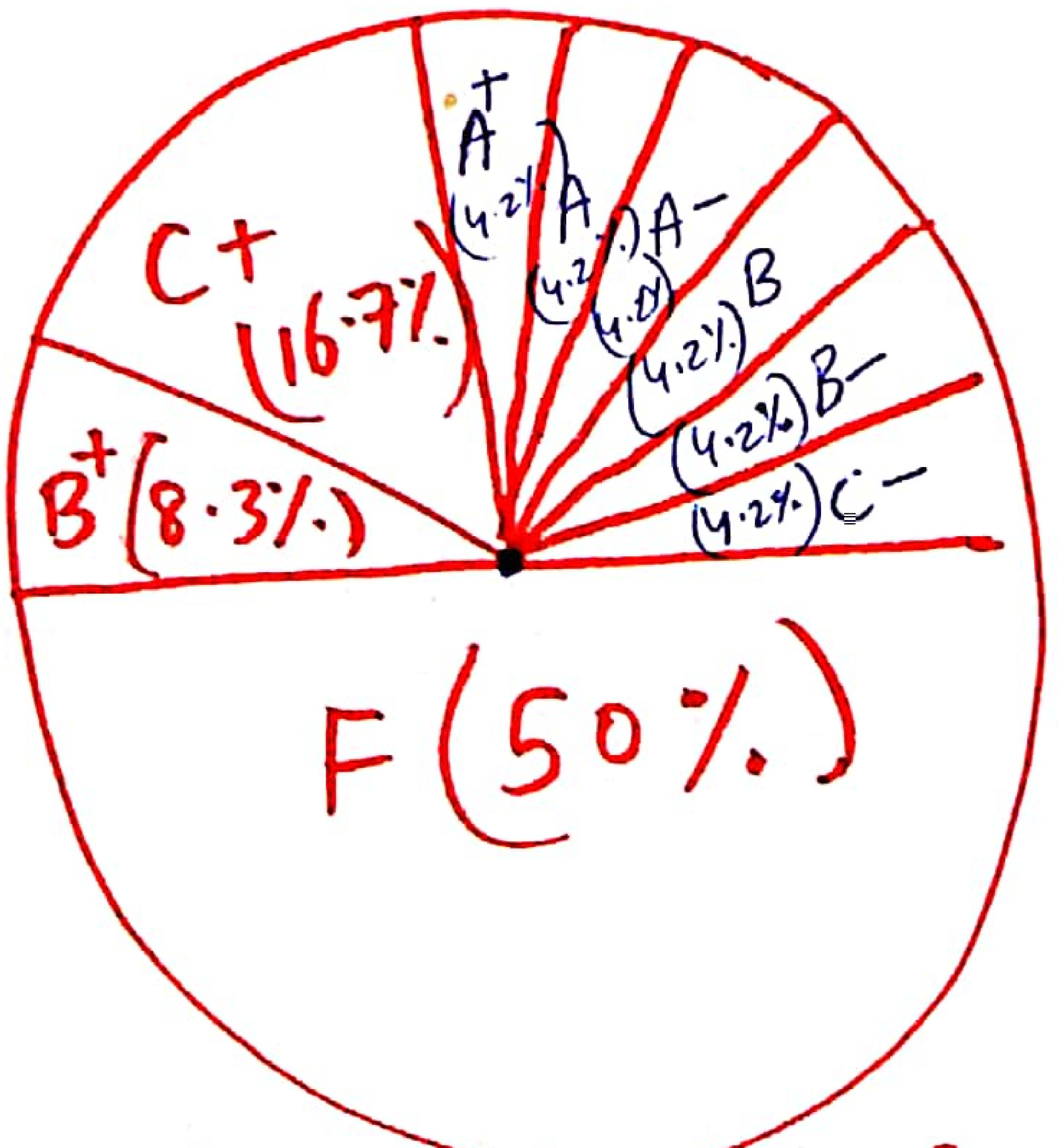
(b)

Total students = 24

1 student has got A^+, 1 student has got A ... and 12 student's have got F grade.

Total grades are 12 $(A^+, A, A-, \ldots, F)$ so we will divide our circle in 12 pieces.

Grade	Count	%
A^+	1	$\frac{1}{24} \times 100 = 4.2$
A	1	4.2
A–	1	4.2
B+	2	8.3
B	1	4.2
B–	1	4.2
C+	4	16.7
C	0	0
C–	1	4.2
D+	0	0
D	0	0
F	12	50



(Rough sketch)

Pie chart of the Grade Report

Sol 3

(a) We know that

interquartile range = Third Quartile − First Quartile

$$(IQR = Q_3 - Q_1)$$

Total number of points $= n = 36$. For Q_1 we have $p = 1/4$ and $np = 36 \cdot \frac{1}{4} = 9$ (integer)

So Q_1 is the average of the values in positions 9 & 10.

So $$Q_1 = \frac{75+77}{2} = \boxed{76}$$

Next for Q_3 we have $P = 3/4$ and $np = 36 \cdot \frac{3}{4}$

$\Rightarrow np = 27$ (integer), so Q_3 is the average of the values in 27^{th} & 28^{th} positions.

So $$Q_3 = \frac{102+107}{2} = \boxed{104.5}$$

Thus

Interquartile Range $= 104.5 - 76 = \boxed{28.5}$

Ans

(b)

The sample standard deviation is best statistic for measuring the variability in a data because it measures variability in linear units.

Scanned with CamScanner

$P(C) = \frac{2}{100}$, $P(H) = \frac{98}{100}$

$P(P_o | C) = \frac{99}{100}$

$P(N | H) = \frac{99.9}{100} \Rightarrow P(P_o | H) = 1 - \frac{99.9}{100}$

$\Rightarrow P(P_o | H) = 0.001$

$P(C | P_o) = ?$

$$P(C | P_o) = \frac{P(P_o | C)\, P(C)}{P(P_o)}$$

$$P(P_o | C) \cdot P(C) = \frac{99}{100} \cdot \frac{2}{100} = 0.0198$$

$$P(P_o) = P(P_o | C) P(C) + P(P_o | H) P(H)$$

$$= \left(\frac{99}{100}\right)\left(\frac{2}{100}\right) + (0.001)\left(\frac{98}{100}\right)$$

$$\Rightarrow P(P_o) = 0.02078$$

so

$$P(C | P_o) = \frac{0.0198}{0.02078} = \underline{\underline{0.95}}$$

$\underline{\underline{95\%}}$

Sol 5

(a) $P\{X \geq 200\} = 1 - P\{X < 200\}$ (Other Method Also possible)

$$= 1 - \int_0^{200} \frac{20000}{(x+100)^3}\,dx$$

$$= 1 - 20000 \int_0^{200} (x+100)^{-3}\,dx$$

$$= 1 - 20000 \left.\frac{(x+100)^{-2}}{-2}\right|_0^{200}$$

$$= 1 + \left.\frac{10000}{(x+100)^2}\right|_0^{200}$$

$$= 1 + \frac{10000}{(200+100)^2} - \frac{10000}{(0+100)^2}$$

$$= 1 + \frac{10000}{90000} - 1$$

$$\Rightarrow \boxed{P\{X \geq 200\} = \frac{1}{9}} \approx \boxed{0.11}$$

(b) $P\{50 < X < 150\} = \int_{50}^{150} \frac{20{,}000}{(x+100)^3}\,dx$

$$= -\left.\frac{10{,}000}{(x+100)^2}\right|_{50}^{150}$$

$$= -\frac{10{,}000}{(150+100)^2} + \frac{10{,}000}{(50+100)^2}$$

$$\boxed{P\{50 < X < 150\} \approx 0.284}$$

Sol 6

Mother's IQ x_i	Daughter's IQ y_i
135	120
127	130
124	114
115	99
104	106
94	92
85	90

$r = ?$ $\quad \bar{x} = 112 , \quad \bar{y} = 107.29$

$\bar{x}\bar{y} = 12016.48$

$\sum_{i=1}^{7} x_i^2 = 89832$ $\quad \sum_{i=1}^{7} y_i^2 = 81897$

$\sum_{i=1}^{7} x_i y_i = 85553$

Thus

$$r = \frac{\sum_{i=1}^{7} x_i y_i - 7\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{7} x_i^2 - 7\bar{x}^2\right)\left(\sum_{i=1}^{7} y_i^2 - 7\bar{y}^2\right)}}$$

$$= \frac{85553 - 7(12016.48)}{\sqrt{\left[89832 - 7(112)^2\right]\left[81897 - 7(107.29^2)\right]}}$$

$\Rightarrow r \approx 0.88$ (Strong Linear relationship b/w Mother & Daughter IQs. If Mother IQ increases then Daug' also