

USE YOUR NOTES TO HELP YOU!!!! You get one page of hand written notes, front and back on the final. Round your values to 3 significant figures, unless otherwise stated.

1) The mean of the ^{twelve} numbers listed below is 6.5.

4, 3, a, 8, 7, 3, 9, 5, 8, 3, 11, 21

~~scribble~~ $\frac{82+a}{12} = 6.5$ $a = -4$

(a) Find the value of a.

(b) Find the median of these numbers.

$med = 6$

2) The table shows the number of children in ⁵⁵ families.

Number of children	Frequency	Cumulative frequency	Rel. freq.
1	3	3	$\frac{3}{55} = .0545$
2	m	22	.345
3	12	34	.218
4	p	q	.164
5	5	48	.0909
6	7	55	.127
	T		

(a) Write down the value of T. 55

(b) Find the values of m, p and q.

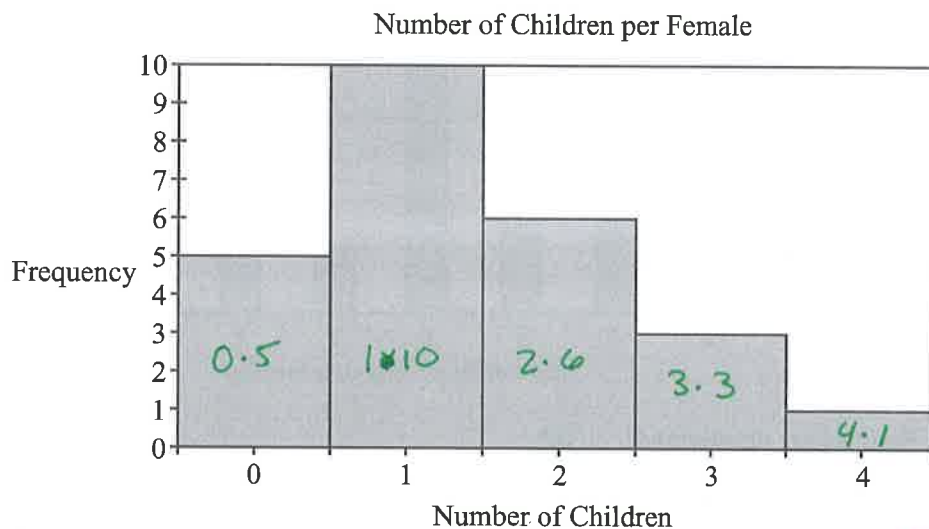
$m = 19$
 $p = 9$
 $q = 43$

(c) Add a relative frequency column

(d) What percent of families have 4 children?

$\approx 16.4\%$

3) A group of 25 females were asked how many children they each had. The results are shown in the histogram below.



(a) Show that the mean number of children per female is 1.4.

(b) Find the mean and standard deviation of the data.

use calculator
for SD

$\frac{0+10+12+9+4}{25} = 1.4$

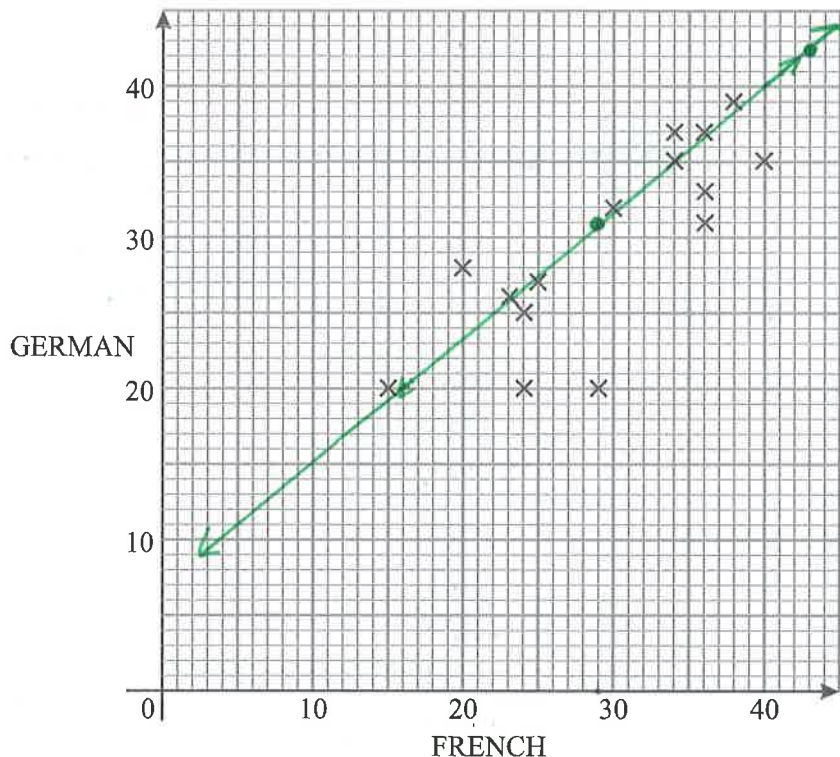
L_1
0
2

L_2
5
10
16
19

1-var.
Freq. list: L_2

$\bar{x} = 1.4$ $\sigma_x = 1.06$

4) The diagram below shows the marks scored by pupils in a French test and a German test. The mean score on the French test is 29 marks and on the German test is 31 marks.



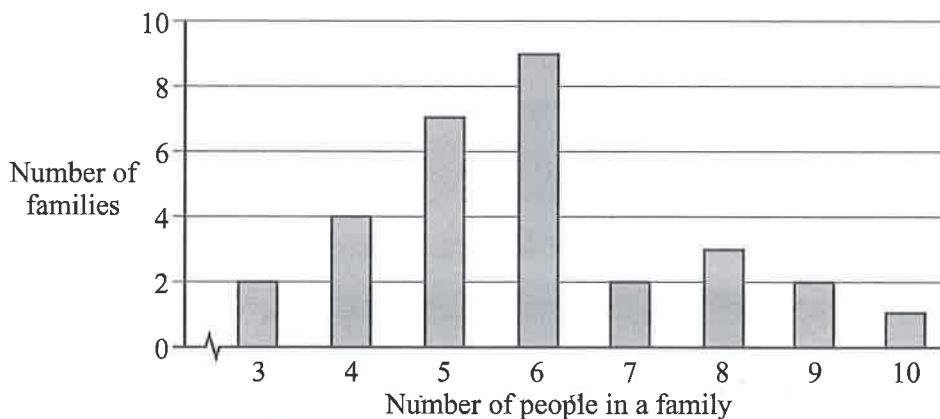
(a) Describe the relationship between the marks scored in the two tests.

a positive, moderate/mod strong correlation.

(b) Draw a line of best fit by eye and extrapolate what a German test mark would be with a French mark of 43.

≈ 42 or 44 marks. [depends on your LOBF]

5) The bar chart below shows the number of people in a selection of families.



(a) How many families are represented? 30

(b) Write down the mode of the distribution. 6

(c) Find, correct to the nearest whole number, the mean number of people in a family.

$$\frac{177}{30} \approx 6 \text{ people}$$

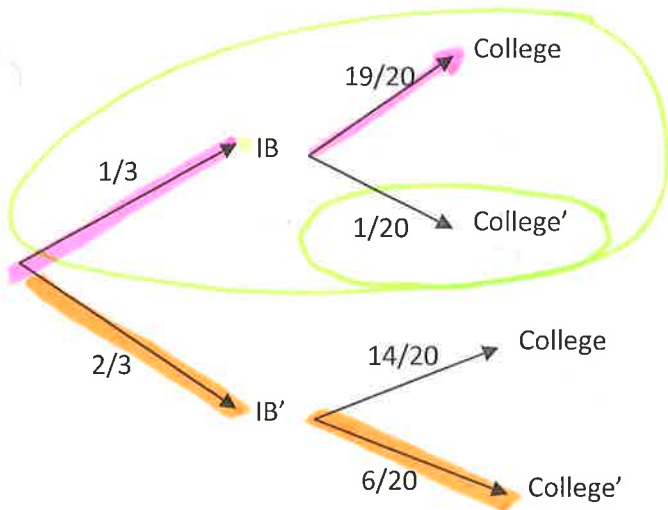
(d) Using a GDC, find the standard deviation.

L1 L2
3 2
4 4
5 7
...

1-var.
freq list: L2

$$\sigma_x \approx 1.72 \text{ people}$$

16. The tree diagram below represents the probability of being an IB student at Creek HS and going to college.



Find:

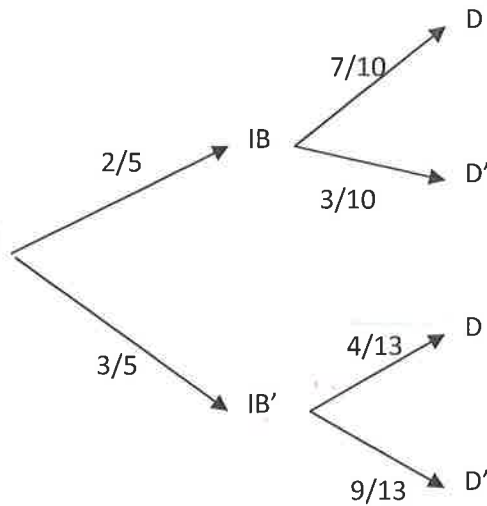
a) $P(\text{the student is in IB and goes to college}) = \frac{1}{3} \cdot \frac{19}{20} = \frac{19}{60}$

b) $P(\text{the student is not in IB and doesn't go to college}) = \frac{2}{3} \cdot \frac{6}{20} = \frac{1}{5}$

c) $P(\text{the student does not go to college given that they are in IB}) = \frac{P(C'|IB)}{P(IB)} = \frac{(1/20)}{(1/3)} = \frac{3}{20}$

17. The tree diagram below represents the probability of living in a dorm at college.

Hint: Use math \blacktriangleright frac on GDC



a) $P(\text{IB or living in a dorm, but not both}) = \left(\frac{3}{10}\right)\left(\frac{2}{5}\right) + \left(\frac{3}{5}\right)\left(\frac{4}{13}\right) = \frac{6}{50} + \frac{12}{65} = \frac{99}{325} \approx .305$

b) $P(\text{student doesn't live in a dorm}) = \left(\frac{2}{5}\right)\left(\frac{3}{10}\right) + \left(\frac{3}{5}\right)\left(\frac{9}{13}\right) = \frac{3}{25} + \frac{27}{65} = \frac{174}{325} \approx .535$

c) $P(\text{student does live in a dorm}) = \left(\frac{2}{5}\right)\left(\frac{7}{10}\right) + \left(\frac{3}{5}\right)\left(\frac{4}{13}\right) = \frac{7}{25} + \frac{12}{65} = \frac{151}{325} \approx .465$

d) $P(\text{IB given that they live in a dorm}) = \frac{P(IB \cap D)}{P(D)} = \frac{7/25}{151/325} = \frac{91}{151} \approx .603$

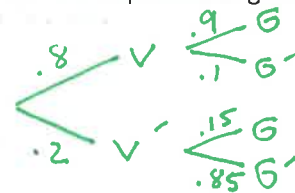
e) $P(\text{not IB given that they don't live in a dorm}) = \frac{P(IB' \cap D')}{P(D')} = \frac{27/65}{174/325} = \frac{45}{58} \approx .776$

f) $P(\text{not IB and went to college and lived in a dorm})$
Not possible.

We do not know who ~~went~~ went to college or not.

18. From experience, it is known that a packet of seeds contains 80% viable seeds. If viable seeds are planted then 90% are expected to grow into plants; and if non-viable seeds are planted then 15% are expected to grow into plants.

a) i) Draw a tree diagram to represent this information.



ii) A farmer sows the seeds from packets containing 1000 seeds. How many plants does he expect to grow?

$$(.8 \times .9) + (.2 \times .15) = .75$$

$$.75 \times 1000 = \boxed{750 \text{ plants}}$$

9. Suppose $P(B) = 0.2$ and $P(A) = 0.4$. Given that A and B are independent events, Draw a Venn diagram and Find:

a. $P(A' \cap B')$

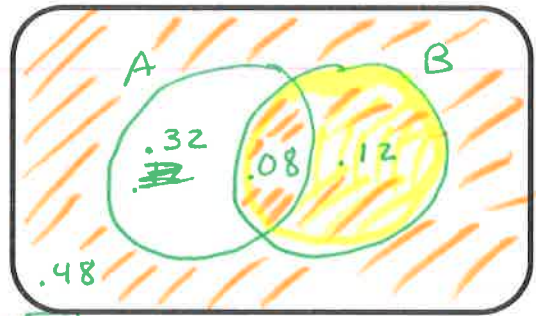
$.2 \times .4 = .08 = P(B \cap A)$

$.48$

b. $P(A' \cup B) = P(A') + P(B) - P(A' \cap B)$

or Shaded = $.68$

c. $P(A'|B) = \frac{P(A' \cap B)}{P(B)}$ or Shaded = $\frac{.12}{.20} = .6$



1. One day the number of customers at three cafés, "Kaylee's Diner" (K), "Sara's Shack" (S) and "Ben's Burgers" (B) was recorded and are given below.

13 were customers of Ben's Burgers only

29 were customers of Sara's Shack only

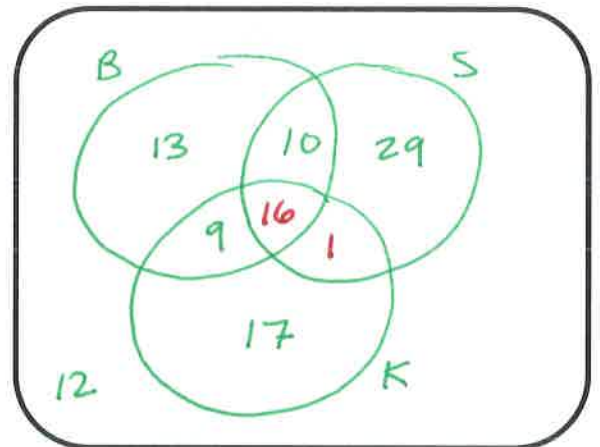
17 were customers of Kaylee's Diner only

10 were customers of Ben's Burgers and Sara's Shack but not Kaylee's Diner

9 were customers of Ben's Burgers and Kaylee's Diner but not Sara's Shack

12 people were records not going into any of the cafés

a. Draw a Venn diagram, using sets labelled K, S and B, that shows this information. Fill in the rest of the diagram as you go through b-e.



There were 48 customers of Ben's Burgers that day.

b. Calculate the number of people who were customers of all three cafés. 16

There were 56 customers of Sara's Shack that day.

c. What was the total number of people?

$n(U) = 107$

d. Find $P(S \cup B) = \frac{13+9+16+10+1+29}{107} = \frac{78}{107}$ or $\approx .729$

e. What is the probability that a randomly selected person (of the universal set) was a customer of Ben's Burgers and Kaylee's diner?

$\frac{25}{107} \approx .234$

f. What is the probability that a randomly selected person was not a customer of Ben's Burgers given that they were a customer from Kaylee's diner?

$P(B'|K) = \frac{P(B' \cap K)}{P(K)} = \frac{N(B' \cap K)}{N(K)} = \frac{18}{43} \approx .419$

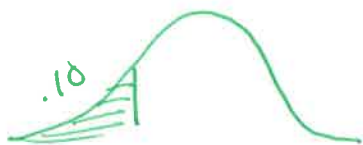
2. You found your calculator! Suppose the times of the horse speeds at the Kentucky Derby are recorded in seconds as follows: $X \sim N(149, 12^2)$.

Find:

a. $P(X \geq 135)$

$$\text{normalcdf}(135, 1E99, 149, 12) \doteq \boxed{.878}$$

b. The slowest time of the fastest 10% of the horse times.



$$\text{invnorm}(.1, 149, 12) \approx \boxed{134 \text{ seconds.}}$$

c. Secretariat ran it in 119.6 seconds. What percent ran SLOWER than him?



$$\text{normalcdf}(119.6, 1E99, 149, 12) \approx .9929$$

$$\approx \boxed{.993\%}$$

d. The Kentucky Derby has had 3000 horses' race total. How many ran slower than 163 seconds?

$$\text{normalcdf}(163, 1E99, 149, 12) \approx .12167 \times 3000 \approx \boxed{365 \text{ horses}}$$

3. The distribution curve shown corresponds to age of NFL players and is:

$$X \sim N(\mu, \sigma^2).$$

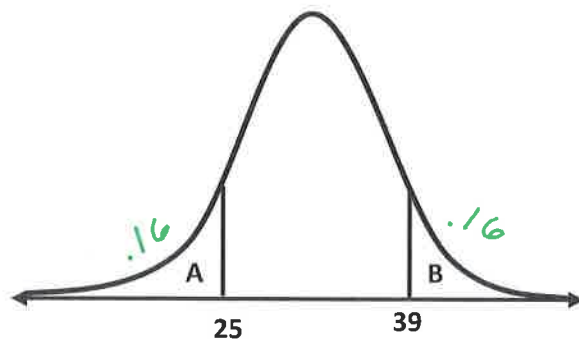
Area A = Area B and $P(A) + P(B) = .32$

a. Find μ and σ (hint: z-scores used)

$$\mu = \boxed{32}$$

$$\text{invnorm}(.16, 0, 1) \approx z = -.99446$$

$$-.99446 = \frac{25 - 32}{\sigma} \quad \sigma \doteq \boxed{7.04}$$



b. Find age k for $P(X > k) = .23$

$$\text{invnorm}(.77, 32, 7.04) \approx \boxed{37.2 \text{ years old}}$$

c. If 500 NFL players were surveyed, how many would you expect to be younger than 20?

$$= \text{normalcdf}(-1E99, 20, 32, 7.04) \times 500 =$$

$$= .04414 \times 500 \doteq \boxed{22.1} \quad \text{or} \quad \boxed{\approx 22 \text{ players}}$$

- 6) Ten students were asked for their average grade at the end of their last year of high school and their average grade at the end of their last year at university. The results were put into a table as follows:

Student	High School grade percent, x	University grade, y
1	90	3.2
2	75	2.6
3	80	3.0
4	70	1.6
5	95	3.8
6	85	3.1
7	90	3.8
8	70	2.8
9	95	3.0
10	85	3.5
Total	835	30.4

- (a) Given that $s_x = 8.96$, $s_y = 0.610$ and $s_{xy} = 4.16$, find the correlation coefficient r , giving your answer to two decimal places. Show your work (without a calculator)

$$\frac{s_{xy}}{s_x s_y} = \frac{4.16}{(8.96)(0.610)} \doteq .761124 \text{ so } \boxed{\approx .76}$$

- (b) Describe the correlation between the high school grades and the university grades.

Strong positive correlation.

- (c) Find the correlation of determination and interpret its value.

$(.761124)^2 \doteq .5793$ or $.579$, so 57.9% of x varies with y .
 Follow the correlation!

- 7) The following shows the type of food a person likes, and their age:

	Mexican	Italian	Indian	American	
<40	44	12	10	39	105
>40	11	31	27	27	96
	55	43	37	66	201

- (a) Calculate the χ^2 for the data shown. How many DF's are there? Use a GDC but show your sums and expected frequency table.

DF: $(2-1)(4-1) = 3$

$\chi^2 \doteq 37.861 \approx 37.9$

	Mex	It.	Ind.	Amer.
<40	28.73	22.46	19.33	34.48
>40	26.27	20.54	17.67	31.52

- (b) If the Critical Value is 7.815, would you reject or accept the null hypotheses: Gender is independent of choice of game. Favorite food.

$37.9 > 7.815$, therefore we reject H_0 , furthermore Gender is not independent of food enjoyment stuff.

9. The following table of observed results gives the number of candidates taking a Mathematics examination classified by gender and grade obtained. Fill in the sums of the table.

		Grade				Total
		6 or 7	4 or 5	3 or 2	1	
Gender	Males	5500	3700	630	23	9853
	Females	6200	4100	1900	9	12209
	Total	11700	7800	2530	32	22062

The question posed is whether gender and grade obtained are independent.

- (a) Fill in the expected contingency table to **Show clearly** that the expected number of males achieving a grade of 6 or 7.

$\frac{11700 \times 9853}{22062}$

		6 or 7	4 or 5	3 or 2	1
Males	5225.3	3483.5	1129.9	14.291	
Females	6474.7	4316.5	1400.1	17.709	

- (b) A χ^2 test is set up.

- (i) State the Null hypothesis.

H_0 : Gender and grade are independent.

- (ii) State the number of degrees of freedom.

$$(2-1)(4-1) = 3$$

- (iii) Calculate the value of χ^2 (GDC is ok).

$\chi^2 = 460$
 $\chi^2 = 459.67$;

Write down the critical value of χ^2 at the 5% level of significance (use notes for this. It will be given for the exam)

$$\chi^2_{crit} = 7.815$$

- (iv) What conclusion is made? Use a rejection inequality

~~45~~ $460 > 7.815$, so we reject H_0 .

- (v) What can you say about gender and grade obtained? Write a sentence.

Gender and grade obtained are not independent.

- (vi) Another test was ran with a different set of data at a 10% significance level. The p-value was .123. What conclusion can be drawn?

.123 > .10, so we do not reject H_0 .

Gender & grade ^{are} independent.

a way different

