FORMULAE AND STATISTICAL TABLES

Standard deviation (sample estimate)

$$\sqrt{\left(\frac{\sum(x-\bar{x})^2}{n-1}\right)}$$

Spearman's rank correlation coefficient

$$1-\frac{6\sum d^2}{n(n^2-1)}$$

Critical values for Spearman's rank

	Level of significance for a one-tailed test				
	0.05	**0.025**	**0.01**	**0.005**	**0.0025**
	Level of significance for a two-tailed test				
N	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
4	1.000	1.000	1.000	1.000	1.000
5	0.700	0.900	0.900	1.000	1.000
6	0.657	0.771	0.829	0.943	0.943
7	0.571	0.679	0.786	0.857	0.893
8	0.548	0.643	0.738	0.810	0.857
9	0.483	0.600	0.683	0.767	0.817
10	0.442	0.564	0.649	0.733	0.782
11	0.418	0.527	0.609	0.700	0.755
12	0.399	0.504	0.587	0.671	0.727
13	0.379	0.478	0.560	0.648	0.698
14	0.367	0.459	0.539	0.622	0.675
15	0.350	0.443	0.518	0.600	0.654
16	0.338	0.427	0.503	0.582	0.632
17	0.327	0.412	0.482	0.558	0.606
18	0.317	0.400	0.468	0.543	0.590
19	0.308	0.389	0.456	0.529	0.575
20	0.299	0.378	0.444	0.516	0.561
21	0.291	0.369	0.433	0.503	0.549
22	0.284	0.360	0.423	0.492	0.537
23	0.277	0.352	0.413	0.482	0.526
24	0.271	0.344	0.404	0.472	0.515
25	0.265	0.337	0.396	0.462	0.505
26	0.260	0.330	0.388	0.453	0.496
27	0.255	0.323	0.381	0.445	0.487
28	0.250	0.317	0.374	0.437	0.479
29	0.245	0.312	0.367	0.430	0.471
30	0.241	0.306	0.361	0.423	0.463

The calculated value must be equal to or exceed the critical value in this table for significance to be shown.

Chi squared distribution formula

$$X^2 = \sum \frac{(O-E)^2}{E} \qquad df = (r-1)(c-1)$$

Critical values for chi-squared distribution

	Level of significance for a one-tailed test					
	0.10	**0.05**	**0.025**	**0.01**	**0.005**	**0.0005**
	Level of significance for a two-tailed test					
df	**0.20**	**0.10**	**0.05**	**0.025**	**0.01**	**0.001**
1	1.64	2.71	3.84	5.02	6.64	10.83
2	3.22	4.61	5.99	7.38	9.21	13.82
3	4.64	6.25	7.82	9.35	11.35	16.27
4	5.99	7.78	9.49	11.14	13.28	18.47
5	7.29	9.24	11.07	12.83	15.09	20.52
6	8.56	10.65	12.59	14.45	16.81	22.46
7	9.80	12.02	14.07	16.01	18.48	24.32
8	11.03	13.36	15.51	17.54	20.09	26.12
9	12.24	14.68	16.92	19.02	21.67	27.88
10	13.44	15.99	18.31	20.48	23.21	29.59
11	14.63	17.28	19.68	21.92	24.73	31.26
12	15.81	18.55	21.03	23.34	26.22	32.91
13	16.99	19.81	22.36	24.74	27.69	34.53
14	18.15	21.06	23.69	26.12	29.14	36.12
15	19.31	22.31	25.00	27.49	30.58	37.70
16	20.47	23.54	26.30	28.85	32.00	39.25
17	21.62	24.77	27.59	30.19	33.41	40.79
18	22.76	25.99	28.87	31.53	34.81	42.31
19	23.90	27.20	30.14	32.85	36.19	43.82
20	25.04	28.41	31.41	34.17	37.57	45.32
21	26.17	29.62	32.67	35.48	38.93	46.80
22	27.30	30.81	33.92	36.78	40.29	48.27
23	28.43	32.01	35.17	38.08	41.64	49.73
24	29.55	33.20	36.42	39.36	42.98	51.18
25	30.68	34.38	37.65	40.65	44.31	52.62
26	31.80	35.56	38.89	41.92	45.64	54.05
27	32.91	36.74	40.11	43.20	46.96	55.48
28	34.03	37.92	41.34	44.46	48.28	56.89
29	35.14	39.09	42.56	45.72	49.59	58.30
30	36.25	40.26	43.77	46.98	50.89	59.70
40	47.27	51.81	55.76	59.34	63.69	73.40
50	58.16	63.17	67.51	71.42	76.15	86.66
60	68.97	74.40	79.08	83.30	88.38	99.61
70	79.72	85.53	90.53	95.02	100.43	112.32

The calculated value must be equal to or less than the critical value in this table for significance to be shown.

Mann-Whitney U test formulae

$$U_a = n_a n_b + \frac{n_a(n_a+1)}{2} - \sum R_a$$

$$U_b = n_a n_b + \frac{n_b(n_b+1)}{2} - \sum R_b$$

(U is the smaller of U_a and U_b)

Critical values for the Mann-Whitney U test

	N_b															
N_a	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
***p* ≤ 0.05 (one-tailed), p ≤ 0.10 (two-tailed)**																
5	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13	15	19	24	28	33	37	42	47	51	56	61	65	70	75	82	84
14	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

	N_b															
N_a	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
***p* ≤ 0.01 (one-tailed), p ≤ 0.02 (two-tailed)**																
5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

	N_b															
N_a	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$p \leq 0.025$ (one-tailed), $p \leq 0.05$ (two-tailed)																
5	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6		5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7			8	10	12	14	16	18	20	22	24	26	28	30	32	34
8				13	15	17	19	22	24	26	29	31	34	36	38	41
9					17	20	23	26	28	31	34	37	39	42	45	48
10						23	26	29	33	36	39	42	45	48	52	55
11							30	33	37	40	44	47	51	55	58	62
12								37	41	45	49	53	57	61	65	69
13									45	50	54	59	63	67	72	76
14										55	59	64	67	74	78	83
15											64	70	75	80	85	90
16												75	81	86	92	98
17													87	93	99	105
18														99	106	112
19															113	119
20																127

	N_b															
N_a	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$p \leq 0.005$ (one-tailed), $p \leq 0.01$ (two-tailed)																
5	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6		2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7			4	6	7	9	10	12	13	15	16	18	19	21	22	24
8				7	9	11	13	15	17	18	20	22	24	26	28	30
9					11	13	16	18	20	22	24	27	29	31	33	36
10						16	18	21	24	26	29	31	34	37	39	42
11							21	24	27	30	33	36	39	42	45	48
12								27	31	34	37	41	44	47	51	54
13									34	38	42	45	49	53	57	60
14										42	46	50	54	48	63	67
15											51	55	60	64	69	73
16												60	65	70	74	79
17													70	75	81	86
18														81	87	92
19															93	99
20																105

The calculated value must be equal to or less than the critical value in this table for significance to be shown.

Wilcoxon Signed Ranks test process

- Calculate the difference between two scores by taking one from the other
- Rank the differences giving the smallest difference Rank 1

  Note: do not rank any differences of 0 and when adding the number of scores, do not count those with a difference of 0, and ignore the signs when calculating the difference
- Add up the ranks for positive differences
- Add up the ranks for negative differences
- T is the figure that is the smallest when the ranks are totalled (may be positive or negative)
- N is the number of scores left, ignore those with 0 difference

Critical values for the Wilcoxon Signed Ranks test

	Level of significance for a one-tailed test		
	0.05	0.025	0.01
	Level of significance for a two-tailed test		
n	0.1	0.05	0.02
N=5	0	-	-
6	2	0	-
7	3	2	0
8	5	3	1
9	8	5	3
10	11	8	5
11	13	10	7
12	17	13	9

The calculated value must be equal to or less than the critical value in this table for significance to be shown.

STATISTICAL KNOWEDGE QUESTIONS

These questions ask you to judge which statistical test is appropriate in which situations. They don't ask you to carry out the test. Here's one from the SAMs (Standardised Assessment Materials) on the Edexcel website.

A group of researchers is testing whether the number of words that can be recalled from a list is affected by age.

One group of participants is under 30 years old, and the other group of participants is over 50 years old.

Participants have to learn and recall words from a list of 100.

Each participant is given a recall score out of 100.

Table 2 shows the data from the investigation.

	Under 30 years old	**Over 50 years old**
Mean score out of 100	22	39

Table 2

(b) Explain which statistical test the researchers could have used to analyse the data.

(2)

.....

.....

.....

.....

The mark scheme explains how marks are awarded like this:

One mark for identifying which statistical test, **related to the data**, should have been used, and one mark for saying that this data is ordinal/interval data and/or a test of difference is being carried out and/or that the study uses an independent groups design/focuses on two ages.

For example:

Mann Whitney/Mann Whitney U test can be used (1) because the data is ordinal/interval data, and it is testing the difference in scores between under 30yrs olds compared to over 50 year olds (1). So it satisfies the conditions.

Answers must relate to the scenario.

Here's another Statistical Knowledge question from the SAMS:

Two psychology students are arguing as to whether males or females would be more likely to stop and help a woman with a baby in a pushchair up the stairs at a busy train station.

They decide to settle the argument by carrying out a structured observation.

(b) State **two** reasons why chi-squared might be an appropriate statistical test for this study.

(2)

.....

.....

.....

.....

Again, the marks are awarded like this:

Two marks for any two of the following points about the study details justifying chi-squared:

- Male versus female helping behaviour is being compared for any significant difference, therefore it needs a test of difference/association (between groups) (1).
- The data gathered is nominal (males or females, help or not). Each person can only belong to one category (1).
- The independent variable is gender (males and females), so the research design is 'independent measures/independent groups'/ 'between subjects design' (1).

STATISTICAL SKILLS QUESTIONS

These questions ask you to carry out the test itself – or part of it. Here's one from the SAMs (Standardised Assessment Materials) on the Edexcel website.

Val was conducting research into aggression. She asked seven participants to rate their own aggression on a scale of 1–9 and then asked the best friend of each participant to give a peer rating of their friend's aggression level using the same scale.

Table 4 shows the data from the investigation.

Complete the table and calculate Spearman's rank correlation coefficient between self-rated aggression and peer-rated aggression.

Self-rated aggression	Rank 1	Peer-rated aggression	Rank 2	d	d^2
2	6.5	3	6		
2	6.5	6	4		
4	5	2	7		
5	4	5	5		
8	3	7	3		
9	1.5	8	2		
9	1.5	9	1		
				Total:	

Table 4

Spearman's rank correlation coefficient .....

4 marks

The mark scheme explains how marks are awarded like this:

One mark for accurate completion of column d (minus signs can be present or not for the mark).

One mark for accurate completion of d^2.

One mark for substituting into equation

$$1 - \frac{6 \times 12}{7(49-1)}$$

One mark for 0.786/0.79.

TRY THESE STATISTICAL QUESTIONS

1. Johannes wants to see if people's concentration gets better after they've eaten. He sets 7 classmates a spot-the-difference test before lunch and counts the number of differences they spot (out of 10). He then sets the same students another, similar test after lunch.

   His results are shown in the table.

	Differences spotted	
	Before lunch	**After lunch**
Charlotte	6	8
Dylan	7	6
James	4	6
Tarion	4	5
Ellen	6	7
Harry	10	8
Megan	9	10

   (a) State the hypothesis Johannes is testing. **(2 marks)**

   (b) Explain **two** reasons why Johannes should use the Wilcoxon statistical test. **(2 marks)**

Johannes calculates that his observed value of W is 8.5.

(c) Using the statistical tables at the front of the book, explain the conclusion Johannes must reach in relation to his hypothesis. **(2 marks)**

2. Delphi has carried out an observation of males and females in the college canteen and whether they choose the vegetarian or the non-vegetarian option. She shows the frequencies in this table:

	Males	Females
Vegetarian	5	9
Non-vegetarian	15	6

(a) What level of data is Delphi recording? **(1 mark)**

Delphi uses the chi squared statistical test.

(b) Complete this table and calculate the value of chi squared for the difference between males and females preferences for vegetarian and non-vegetarian meals. **(4 marks)**

	A	B	C	D
O	5	15	9	6
E	8	12	6	9
O - E				
$(O - E)^2$				
$\frac{(O - E)^2}{E}$				

Chi Squared …….

3. Gloria wonders why so many of her friends who watch the violent TV drama *Throne of Bones* also enjoy ice hockey. She decides to investigate whether the people who like *Throne of Bones* the most are also the biggest ice hockey fans. She does this by asking 12 friends to rate their liking for *Throne of Bones* on a 1-10 scale and also rate how much they enjoy ice hockey, also on a 1-10 scale.

   (a) Outline a 1-tailed hypothesis for Gloria's investigation. **(2 marks)**

   (b) What sort of graph would best express her results? **(1 mark)**

   (c) Which statistical test should Gloria use to interpret her results? **(1 mark)**

   (d) Explain **two** reasons why the test you identified in (c) would be appropriate for Gloria's hypothesis. **(2 marks)**

   When Gloria carries out her test, she arrives at a calculated value of 0.53.

   (e) Using the tables at the start of the booklet, write a statement of significance for Gloria's data in relation to her hypothesis. **(2 marks)**

4. Greta and Paige are arguing about the best way to revise for the test. Greta likes spaced revision, revising small amounts over a long period of time. Paige prefers to 'cram' right before the test.

   After the test, Greta asks their 12 classmates how they revised: there were 7 'crammers' and 5 who did spaced revision. Greta thinks the results show that spaced revision is best.

   (a) Write a null hypothesis for Greta's research. **(2 marks)**

This table shows the grades the class got and how Greta ranked them.

Cramming (Condition A)	**Rank (A)**	**Spaced Revision (Condition B)**	**Rank (B)**
A	2.5	A	2.5
D	8.5	B	4.5
E	11.5	D	8.5
D	8.5	A*	1
D	8.5	B	4.5
E	11.5		
C	6		

(b) What grades were the modes in each condition? **(2 marks)**

Condition A: Condition B:

Greta starts to carry out a Mann-Whitney U Test on the results,

(c) Explain why the Mann-Whitney U Test is the most appropriate statistical test for Greta to use. **(2 marks)**

Greta works out the calculated value of U_b for Condition B (Spaced Revision) and it is 29. She then compares it to the critical value for a 1-tailed test at a probability of $p \leq 0.05$.

(d) Identify the critical value. **(1 mark)**

(e) Calculate the value of U_a for Condition A (Cramming) using the formula at the start of the book. **(2 marks)**

U_a = ……..

(f) Explain whether Greta should reject her null hypothesis. **(2 marks)**