UNIVERSITY OF SWAZILAND

FINAL EXAMINATION PAPER 2015

TITLE OF PAPER	:	INFERENTIAL STATISTICS
COURSE CODE	:	ST 220
TIME ALLOWED	:	TWO (2) HOURS
REQUIREMENTS	:	CALCULATOR AND STATISTICAL TABLES
INSTRUCTIONS	:	THIS PAPER HAS SIX (6). ANSWER ANY THREE (3) QUESTIONS.

Question 1 [20 marks, 12+8]

(a) One criterion that students may apply when choosing a university for undergraduate studies is the number of scheduled contact hours per week. However, different disciplines have developed teaching patterns best suited to their subject material. Typically, scientific and engineering subjects may require more contact hours to allow for structured programmes of lectures and laboratory sessions. In contrast, arts and humanities subjects require fewer contact hours to foster independent, creative and original analysis.

The table below shows the sample mean numbers of contact hours for ten discipline categories at two universities, A and B. The values have been obtained from samples of students in each discipline at each university. It is of interest to investigate whether there is a significant overall difference between the mean numbers of contact hours at these two universities.

Subject category	Mean number of hours at university A	Mean number of hours at university B
Medical Studies	20.1	20.3
Technology and Engineering	19.2	21.4
Law	11.6	12.3
Business and Management Studies	12.2	13.1
Philosophical and Historical Studies	8.2	7.9
Classics and Linguistics	10.1	10.7
Mathematics and IT	14.9	16.1
Physical and Biological Sciences	16.0	17.8
Urban, Regional and Architectural Studies	16.2	15.8
Media and Social Studies	12.0	12.8
TOTAL	140.5	148.2

Perform a t test at the 5% significance level to examine whether the difference in contact hours between the universities has a mean of zero. State your null and alternative hypotheses clearly and report your conclusions. State any assumptions made in carrying out the t test.

(b) In 2001, the Supreme Court, by a vote of 8-0, struck down state laws that legalized marijuana for medicinal purposes. The Gallup Organization later conducted surveys of randomly selected individuals (18+ years) and asked them whether they support the limited use of marijuana when prescribed by physicians to relive pain and suffering. The results of the survey by age group, are as follows:

	Age		
Opinion	**18-29**	**30-49**	**60+**
For	172	313	258
Against	52	103	119

Is there evidence to indicate that the proportions of individuals in each age group who are for the legalization of marijuana for medicinal use if different at the $\alpha = 0.01$ level of significance.

Question 2 [20 marks, 9+8+3]

A city council is considering introducing a congestion charge for motorists travelling into or out of the city centre. The city is divided into ten administrative areas. In order to assess the popularity of such a measure, samples of residents from two of the administrative areas are asked whether or not they are in favour of the introduction of the congestion charge. The results are shown below.

	In favour of the charge	Not in favour of the charge
Area 1	61	95
Area 2	20	84

(a) Perform a χ^2 test at the 5% significance level to investigate whether there is an association between the area of the city and the attitude to the proposed congestion charge. State your null hypothesis and report your conclusions.

(b) Estimate the proportions who are not in favour of the proposed congestion charge for each of the two areas and calculate an approximate 95% confidence interval for the difference in these two proportions.

(c) You could perform a hypothesis test to examine whether there is a difference in the proportions of those who are not in favour of the proposed congestion charge in the two areas. Without performing this test, outline briefly how its results would relate to your answers to parts (a) and (b).

Question 3 [20 marks, 6+8+1+5]

(b) The amount of a potentially toxic pollutant in the water of a river affects the edibility of mussels grown in its estuary. An environmental health officer has heard a report of a leak of this pollutant into the estuary and undertakes an investigation into how this has affected the mussel population. He takes a sample of ten mussels randomly from this population and measures the amount of the pollutant in parts per million (ppm) in each of them. These ten values are as follows.

```
39.5   38.6   44.9   36.4   45.6   46.6   36.1   32.3   35.0   35.5
```

(i) Calculate a 99% confidence interval for the population mean.

Official health guidelines state that mussels are safe to eat provided that the (population) mean level of pollutant does not exceed 36 ppm.

(i) Test, at the 1% significance level, whether or not the population mean level exceeds 36 ppm. State the null and alternative hypotheses and report your conclusions.

(ii) State briefly why the 99% confidence interval for the mean calculated in part (i) cannot be used directly to perform the required hypothesis test in part (iii).

(b) A blended wine is intended to comprise two parts of Sauvignon to one part of Merlot. The amounts dispensed to make up a nominal 75cl bottle of this wine are X cl of Sauvignon and Y cl of Merlot, where X and Y are assumed to be independent Normally distributed random variables with respective means 52 and 26 cl and respective variances 1 and 0.5625. Find the probability that the actual volume of wine dispensed into a bottle is less than the nominal volume.

Question 4 **[20 marks, 8+8+4]**

(a) A manufacturer of luxury cosmetics has recently put a new product on the market. This product is initially being offered at a wide range of prices, and the company has made a survey of its sales y (in 100s) and prices x (in £) across a random sample of stores in which it is sold. It wishes to examine whether, on the whole, increased price is associated with reduced sales. The results are shown in the following table.

Store	1	2	3	4	5	6	7	8	9	10
Price x (£)	27	30	37	47	55	62	70	80	95	99
Sales y (100s)	110	79	69	48	51	44	29	32	26	30

A research assistant suggests calculating the product-moment correlation coefficient, r, between sales and prices. Carry out this calculation and test at the 1% significance level the null hypothesis of zero correlation against an appropriate one-sided alternative. You are given that

$$\sum x = 602, \quad \sum x^2 = 42202, \quad \sum y = 518, \quad \sum y^2 = 33384, \quad \sum xy = 25712.$$

(b) The vitamin content of the flesh of each of a random sample of eight oranges and of a random sample of five lemons was measured. The results are given in milligrams per 10 grams.

Oranges	1.14	1.59	1.57	1.33	1.08	1.27	1.43	1.36
Lemons	1.04	0.95	0.63	1.62	1.11			

(i) Estimate a 99% confidence interval for the difference between the mean vitamin contents. Based on the confidence interval would you conclude that the difference between the two means is significantly different from zero?

(ii) What assumption(s) were made in the computation of the above confidence interval?

Question 5 **[20 marks, 4+8+4+4]**

(a) State the model and assumptions for the one-way analysis of variance, defining your notation.

(b) A commute in a large city can travel to work by car, bicycle or bus. She times four journeys by each method with the following results, in minutes.

Car	Bicycle	Bus
27	34	26
45	38	41
33	43	35
31	42	46

Carry out an analysis of variance and test at the 5% significance level whether there are differences in the mean journey times for the three methods of transport.

(c) A certain brand of beans is sold in tins, the tins being filled and sealed by a machine. the mass of beans in each tin is normally distributed with mean 425 g and a standard deviation of 25 g and the mass of the tin is normally distributed with mean 90 g and standard deviation 10 g.

Find the probabilities that the total mass of the sealed tin and its beans

(i) exceeds 550 g,

(ii) lies between 466 g and 575 g.

Question 6 [20 marks, 2+8+4+3+3]

100 men are surveyed as to whether they play cricket, tennis or golf. It is found that

- 10 play none of these sports
- 5 play all three of these sports
- 88 play cricket or tennis or both
- 78 play cricket or golf or both
- 30 play golf and tennis but not cricket
- 38 play golf
- 74 play tennis.

Find the following.

(a) The number of the men who play at least one of these sports.

(b) The number of the men who play exactly one of these sports.

(c) The number of the men who play exactly two of these sports.

(d) Of those who do not play golf, the proportion who play cricket.

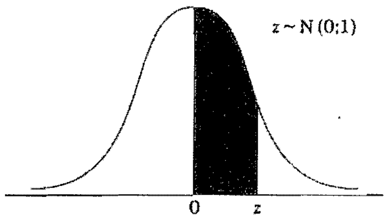
(e) The mean number of sports played by these men.

APPENDIX 1: LIST OF STATISTICAL TABLES

TABLE 1

The standard normal distribution (z)

This table gives the area under the standard normal curve between 0 and z i.e. $P[0 < Z < z]$

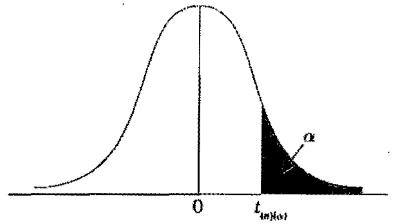


Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3.0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900
3.1	0.49903	0.49906	0.49910	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3.2	0.49931	0.49934	0.49936	0.49938	0.49940	0.49942	0.49944	0.49946	0.49948	0.49950
3.3	0.49952	0.49953	0.49955	0.49957	0.49958	0.49960	0.49961	0.49962	0.49964	0.49965
3.4	0.49966	0.49968	0.49969	0.49970	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3.5	0.49977	0.49978	0.49978	0.49979	0.49980	0.49981	0.49981	0.49982	0.49983	0.49983
3.6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3.7	0.49989	0.49990	0.49990	0.49990	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3.8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997
4.0	0.49997	0.49997	0.49997	0.49997	0.49997	0.49997	0.49998	0.49998	0.49998	0.49998

TABLE 2

The t distribution

This table gives the value of $t_{(n)(\alpha)}$ where n is the degrees of freedom i.e. ■ $= P[t \geq t_{(n)(\alpha)}]$

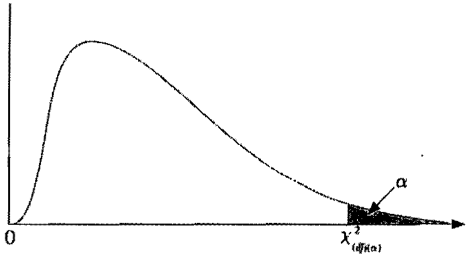


n	0.100	0.050	0.025	0.010	0.005	0.0025
1	3.078	6.314	12.706	31.821	63.657	127.321
2	1.886	2.920	4.303	6.965	9.925	14.089
3	1.638	2.353	3.182	4.541	5.841	7.453
4	1.533	2.132	2.776	3.747	4.604	5.598
5	1.476	2.015	2.571	3.365	4.032	4.773
6	1.440	1.943	2.447	3.143	3.707	4.317
7	1.415	1.895	2.365	2.998	3.499	4.029
8	1.397	1.860	2.306	2.896	3.355	3.833
9	1.383	1.833	2.262	2.821	3.250	3.690
10	1.372	1.812	2.228	2.764	3.169	3.581
11	1.363	1.796	2.201	2.718	3.106	3.497
12	1.356	1.782	2.179	2.681	3.055	3.428
13	1.350	1.771	2.160	2.650	3.012	3.372
14	1.345	1.761	2.145	2.624	2.977	3.326
15	1.341	1.753	2.131	2.602	2.947	3.286
16	1.337	1.746	2.120	2.583	2.921	3.252
17	1.333	1.740	2.110	2.567	2.898	3.222
18	1.330	1.734	2.101	2.552	2.878	3.197
19	1.328	1.729	2.093	2.539	2.861	3.174
20	1.325	1.725	2.086	2.528	2.845	3.153
21	1.323	1.721	2.080	2.518	2.831	3.135
22	1.321	1.717	2.074	2.508	2.819	3.119
23	1.319	1.714	2.069	2.500	2.807	3.104
24	1.318	1.711	2.064	2.492	2.797	3.091
25	1.316	1.708	2.060	2.485	2.787	3.078
26	1.315	1.706	2.056	2.479	2.779	3.067
27	1.314	1.703	2.052	2.473	2.771	3.057
28	1.313	1.701	2.048	2.467	2.763	3.047
29	1.311	1.699	2.045	2.462	2.756	3.038
30	1.310	1.697	2.042	2.457	2.750	3.030
31	1.309	1.696	2.040	2.453	2.744	3.022
32	1.309	1.694	2.037	2.449	2.738	3.015
33	1.308	1.692	2.035	2.445	2.733	3.008
34	1.307	1.691	2.032	2.441	2.728	3.002
35	1.306	1.690	2.030	2.438	2.724	2.996
36	1.306	1.688	2.028	2.434	2.719	2.990
37	1.305	1.687	2.026	2.431	2.715	2.985
38	1.304	1.686	2.024	2.429	2.712	2.980
39	1.304	1.685	2.023	2.426	2.708	2.976
40	1.303	1.684	2.021	2.423	2.704	2.971
45	1.301	1.679	2.014	2.412	2.690	2.952
50	1.299	1.676	2.009	2.403	2.678	2.937
60	1.296	1.671	2.000	2.390	2.660	2.915
70	1.294	1.667	1.994	2.381	2.648	2.899
80	1.292	1.664	1.990	2.374	2.639	2.887
90	1.291	1.662	1.987	2.368	2.632	2.878
100	1.290	1.660	1.984	2.364	2.626	2.871
110	1.289	1.659	1.982	2.361	2.621	2.865
120	1.289	1.658	1.980	2.358	2.617	2.860
140	1.288	1.656	1.977	2.353	2.611	2.852
160	1.287	1.654	1.975	2.350	2.607	2.847
180	1.286	1.653	1.973	2.347	2.603	2.842
200	1.286	1.653	1.972	2.345	2.601	2.839
∞	1.282	1.645	1.960	2.326	2.576	2.807

TABLE 3

The Chi-Squared distribution (χ^2)

This table gives the value of $\chi^2_{(df)(\alpha)}$ where df is the degrees of freedom i.e. ■ = $P[\chi^2 > \chi^2_{(df)(\alpha)}]$

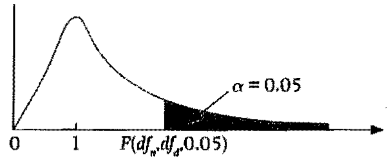


df \ α	0.100	0.050	0.025	0.01	0.005	0.0025
1	2.707	3.843	5.026	6.635	7.880	9.142
2	4.605	5.992	7.378	9.210	10.597	11.983
3	6.251	7.815	9.348	11.345	12.838	14.321
4	7.779	9.488	11.143	13.277	14.860	16.424
5	9.236	11.071	12.833	15.086	16.750	18.386
6	10.645	12.592	14.449	16.812	18.548	20.249
7	12.017	14.067	16.013	18.475	20.278	22.040
8	13.362	15.507	17.535	20.090	21.955	23.774
9	14.684	16.919	19.023	21.666	23.589	25.462
10	15.987	18.307	20.483	23.209	25.188	27.112
11	17.275	19.675	21.920	24.725	26.757	28.729
12	18.549	21.026	23.337	26.217	28.300	30.318
13	19.812	22.362	24.736	27.688	29.819	31.883
14	21.064	23.685	26.119	29.141	31.319	33.426
15	22.307	24.996	27.488	30.578	32.801	34.950
16	23.542	26.296	28.845	32.000	34.267	36.456
17	24.769	27.587	30.191	33.409	35.718	37.946
18	25.989	28.869	31.526	34.805	37.156	39.422
19	27.204	30.144	32.852	36.191	38.582	40.885
20	28.412	31.410	34.170	37.566	39.997	42.336
21	29.615	32.671	35.479	38.932	41.401	43.775
22	30.813	33.924	36.781	40.289	42.796	45.204
23	32.007	35.172	38.076	41.638	44.181	46.623
24	33.196	36.415	39.364	42.980	45.558	48.034
25	34.382	37.652	40.646	44.314	46.928	49.435
26	35.563	38.885	41.923	45.642	48.290	50.829
27	36.741	40.113	43.195	46.963	49.645	52.215
28	37.916	41.337	44.461	48.278	50.994	53.594
29	39.087	42.557	45.722	49.588	52.335	54.967
30	40.256	43.773	46.979	50.892	53.672	56.332
31	41.422	44.985	48.232	52.191	55.003	57.692
32	42.585	46.194	49.480	53.486	56.328	59.045
33	43.745	47.400	50.725	54.776	57.648	60.395
34	44.903	48.602	51.966	56.061	58.964	61.738
35	46.059	49.802	53.203	57.342	60.275	63.076
36	47.212	50.998	54.437	58.619	61.581	64.410
37	48.363	52.192	55.668	59.892	62.883	65.739
38	49.513	53.384	56.896	61.162	64.181	67.063
39	50.660	54.572	58.120	62.428	65.476	68.383
40	51.805	55.758	59.342	63.691	66.766	69.699
45	57.505	61.656	65.410	69.957	73.166	76.223
50	63.167	67.505	71.420	76.154	79.490	82.664
60	74.397	79.082	83.298	88.379	91.952	95.344
70	85.527	90.531	95.023	100.425	104.215	107.808
80	96.578	101.879	106.629	112.329	116.321	120.102
90	107.565	113.145	118.136	124.116	128.299	132.252
100	118.498	124.342	129.561	135.807	140.169	144.293
110	129.385	135.480	140.917	147.414	151.948	156.271
120	140.233	146.567	152.211	158.950	163.648	168.154
140	161.827	168.613	174.648	181.840	186.847	191.759
[illegible]	190.522	196.525	202.786	209.252	214.340	218.508
[illegible]	212.310	219.856	225.296	232.647	237.880	242.305

TABLE 4 (a)

F distribution ($\alpha = 0.05$)

The entries in this table are critical values of F for which the area under the curve to the right is equal to 0.05.



Degrees of freedom for denominator	Degrees of freedom for numerator 1	2	3	4	5	6	7	8	9	10
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

TABLE 4 (a) continued

F distribution ($\alpha = 0.05$)

	Degrees of freedom for numerator								
Degrees of freedom for denominator	12	15	20	24	30	40	60	120	∞
1	243.9	245.9	248	249.1	250.1	251.1	252.2	253.3	254.3
2	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

APPENDIX 2: LIST OF KEY FORMULAE

MEASURES OF CENTRAL LOCATION

Arithmetic mean — *Ungrouped data*

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad 3.1$$

Grouped data

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{n}$$

Mode — *Grouped data*

$$M_o = O_{mo} + \frac{c(f_m - f_{m-1})}{2f_m - f_{m-1} - f_{m+1}} \qquad 3.3$$

Median — *Grouped data*

$$M_e = O_{me} + \frac{c\left[\frac{n}{2} - f(<)\right]}{f_{me}} \qquad 3.2$$

Lower quartile — *Grouped data*

$$Q_1 = O_{q1} + \frac{c\left(\frac{n}{4} - f(<)\right)}{f_{q1}} \qquad 3.7$$

Upper quartile — *Grouped data*

$$Q_3 = O_{q3} + \frac{c\left(\frac{3n}{4} - f(<)\right)}{f_{q3}} \qquad 3.8$$

Geometric mean — *Ungrouped data*

$$GM = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n} \qquad 3.4$$

Weighted arithmetic mean — *Grouped data*

$$\text{weighted } \bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} \qquad 3.5$$

MEASURES OF DISPERSION AND SKEWNESS

Range

Range = Maximum value – Minimum value + 1
$= x_{max} - x_{min} + 1$ 3.9

Variance

Mathematical – ungrouped data

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{(n-1)}$$ 3.10

Computational – ungrouped data

$$s^2 = \frac{\Sigma x_i^2 - n\bar{x}^2}{(n-1)}$$ 3.11

Standard deviation

$$s = \sqrt{s^2}$$ 3.12

Coefficient of variation

$$CV = \frac{s}{\bar{x}} \times 100\%$$ 3.13

Pearson's coefficient of skewness

$$sk_p = \frac{n\Sigma(x_i - \bar{x})^3}{(n-1)(n-2)s^3}$$ 3.14

$$sk_p = \frac{3\text{ (Mean – Median)}}{\text{Standard deviation}}$$ (approximation) 3.15

PROBABILITY CONCEPTS

Conditional probability

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$ 4.2

Addition rule

Non-mutually exclusive events
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ 4.3

Mutually exclusive events
$P(A \cup B) = P(A) + P(B)$ 4.4

Multiplication rule

Statistically dependent events
$P(A \cap B) = P(A/B) \times P(B)$ 4.5

Statistically independent events
$P(A \cap B) = P(A) \times P(B)$ 4.6

***n*! = *n* factorial**

$n \times (n-1) \times (n-2) \times (n-3) \times \ldots \times 3 \times 2 \times 1$ 4.8

Permutations

$${}_nP_r = \frac{n!}{(n-r)!}$$ 4.10

Combinations

$${}_nC_r = \frac{n!}{r!\,(n-r)!}$$ 4.11

PROBABILITY DISTRIBUTIONS

Binomial distribution

$P(x) = {}_nC_x\, p^x(1-p)^{(n-x)}$ for $x = 0, 1, 2, 3, \ldots, n$ 5.1

$P(x \text{ successes}) = \frac{n!}{x!\,(n-x)!} p^x(1-p)^{(n-x)}$ for $x = 0, 1, 2, 3, \ldots, n$

Binomial descriptive measures

Mean $\mu = np$
Standard deviation $\sigma = \sqrt{np(1-p)}$ 5.2

Poisson distribution

$P(x) = \frac{e^{-a} a^x}{x!}$ for $x = 0, 1, 2, 3 \ldots$ 5.3

Poisson descriptive measures

Mean $\mu = a$
Standard deviation $\sigma = \sqrt{a}$ 5.4

Standard normal probability

$$z = \frac{x - \mu}{\sigma}$$ 5.6

CONFIDENCE INTERVALS

Single mean — *n large; variance known*

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z\frac{\sigma}{\sqrt{n}} \quad 7.1$$

(lower limit) (upper limit)

n small; variance unknown

$$\bar{x} - t_{(n-1)}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{(n-1)}\frac{s}{\sqrt{n}} \quad 7.2$$

(lower limit) (upper limit)

Single proportion

$$p - z\sqrt{\frac{p(1-p)}{n}} \le \pi \le p + z\sqrt{\frac{p(1-p)}{n}} \quad 7.3$$

(lower limit) (upper limit)

HYPOTHESES TESTS

Single mean — *Variance known*

$$z\text{-}stat = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad 8.1$$

Variance unknown; n small

$$t\text{-}stat = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \quad 8.2$$

Single proportion

$$t\text{-}stat = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \quad 8.3$$

Difference between two means — *Variances known*

$$z\text{-}stat = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad 9.1$$

Variances unknown; n_1 and n_2 small

$$t\text{-}stat = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad 9.2$$

Paired *t*-test

$$t\text{-}stat = \frac{\bar{x}_d - \mu_d}{\frac{s_d}{\sqrt{n}}} \quad 9.5$$

where $\mu_d = (\mu_1 - \mu_2)$

and $s_d = \sqrt{\frac{\Sigma(x_d - \bar{x}_d)^2}{n - 1}}$

Differences between two proportions

$$z\text{-}stat = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{\pi}(1 - \hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where } \hat{\pi} = \frac{x_1 + x_2}{n_1 + n_2};\ p_1 = \frac{x_1}{n_1};\ p_2 = \frac{x_2}{n_2} \quad 9.8$$

Chi-Squared

$$\chi^2\text{-}stat = \Sigma\frac{(f_o - f_e)^2}{f_e} \quad 10.1$$

Overall mean

$$\bar{\bar{x}} = \frac{\Sigma\Sigma x_{ij}}{N} \quad 11.2$$

Total sum of squares (SSTotal) $= \sum_i \sum_j (x_{ij} - \bar{\bar{x}})^2$ 11.3

SST $= \sum_j^k n_j(\bar{x}_j - \bar{\bar{x}})^2$ 11.4

SSE $= \sum_j \sum_i (x_{ij} - \bar{x}_j)^2$ 11.5

SSTotal = SST + SSE 11.6

MSTotal $= \frac{\text{SSTotal}}{N - 1}$ 11.7

MST $= \frac{\text{SST}}{k - 1}$ 11.8

MSE $= \frac{\text{SSE}}{N - k}$ 11.9

***F*-stat** $= \frac{\text{MST}}{\text{MSE}}$ 11.10