UNIVERSITY OF SWAZILAND

SUPPLEMENTARY EXAMINATION PAPER 2013

TITLE OF PAPER	:	INFERENTIAL STATISTICS
COURSE CODE	:	ST 220
TIME ALLOWED	:	TWO (2) HOURS
REQUIREMENTS	:	CALCULATOR AND STATISTICAL TABLES
INSTRUCTIONS	:	THIS PAPER HAS FIVE (5) QUESTIONS. ANSWER ANY THREE (3) QUESTIONS.

Question 1 [20 marks, 10+7+3]

(a) Random samples are taken from two populations with distributions $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$ (i.e. their variances are the same). The summary statistics for the two samples are shown in the following Table:

	Sample Size n	**Sample Mean** m	**Sample Variance** s^2
x-data	19	7.0	1.69
y-data	25	5.1	2.56

Compute a 95% confidence interval for the difference $\mu_X - \mu_Y$ between the two population means. Does the result support the view that there is no true difference between the population means? (Explain your reasoning!)

(b) A market research company has conducted a survey of adults in two large towns, either side of an international border, in order to judge attitudes towards a controversial internationally broadcast celebrity television programme. The following table shows some of the information obtained by the survey:

	Town A	**Town Z**
Sample size	40	40
Sample number approving of the programme	24	22

(i) Conduct a formal hypothesis test, at the 5% significance level, of the claim that the population proportions approving the programme in the two towns are equal.

(ii) Would your conclusion be the same if, in both towns the sample sizes had been 100 (with the same sample proportions of approvals)?

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

(a) After a recent National Aids Awareness Campaign, a market research company conducted a countrywide survey on behalf of the Department of National Health. The brief was to establish whether the *recall rate* of *teenagers* differed from that of *young adults* between 20 and 30 years of age.

The market research company interviewed 640 teenagers and 420 young adults countrywide. Three hundred and sixty-two teenagers recalled the Aids Awareness slogan used during the campaign, and 260 young adults were able to recall the same Aids Awareness slogan of "Aids: don't let it happen".

Test, at the 5% level of significance, the hypothesis that there is an *equal recall rate* between teenagers and young adults.

(b) An aircraft maintenance company bought equipment for detecting structural defects in aircrafts. Tests indicate that 95% of the time the equipment detects defects when they actually exist, and 1% of the time it gives a false alarm that indicates the presence of a structural defect when in fact there is none. If 2% of the aircrafts actually have structural defects, what is the probability that an aircraft actually has a structural defect given that the equipment indicates that it has a structural defect?

(c) If $\sigma_X = 1$, determine the number of observations required to ensure that at the 99% confidence level, $\bar{X} - 0.1 \leq \mu \leq \bar{X} + 0.1$ where

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

(d) A random sample of size 81 is taken from a population that has a mean of 24 and variance 324. Use the central limit theorem to determine the probability that the sample mean lies between 23.9 and 24.2.

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

(a) Researchers in the USA investigated the effects of alcohol consumption on brain size. A sample of 10 non-drinkers and 8 heavy drinkers (more than 14 units of alcohol per week) was obtained and each person in the sample had a brain scan. The ratios of brain volume to skull size obtained for the eighteen people in the sample are given in the table.

Non-drinkers	0.794	0.798	0.799	0.802	0.803	0.804	0.805	0.806	0.809	0.810
Heavy drinkers	0.785	0.787	0.789	0.791	0.792	0.796	0.797	0.801		

Assuming that the sample is random, carry out a test, at the 5% level of significance, to investigate the claim that heavy drinkers have a smaller average ratio of brain volume to skull size than non-drinkers. Interpret your conclusion in context.

(b) In a diet test, each of four diet programs is applied to a sample of people. At the end of three weeks, the amount of pounds people lost are shown below.

Diet Program			
1	2	3	4
12	19	16	28
6	10	20	17
18	13	26	22
23	20	19	16
	25		20

Test to determine if there is enough evidence at the 5% significance level to infer that at least two population locations differ. State the hypothesis, critical region(s) and conclusions. Show all calculations.

(c) 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, 6+4+6+4]**

(a) Music Technologies, an electronics retail company in Durban has kept records of the number of ipods sold within a week of placing advertisements in the *Mercury*. The following table shows the *number of ipods sold* and the corresponding *number of advertisements placed* in the *Mercury* for 12 randomly selected weeks over the past year.

Ads	4	4	3	2	5	2	4	3	5	5	3	4
Sales	26	28	24	18	35	24	36	25	31	37	30	32

(i) Estimate the linear regression line ($\sum x^2 = 174$, $\sum xy = 1324$ and $\sum y^2 = 10336$).

(ii) Compute and interpret the coefficient of determination.

(iii) Is the *relationship* between the number of *newspaper advertisements* placed and *ipod sales* meaningful (or significant)? Use $\alpha = 0.05$.

(b) Every morning, Duncan bakes 30 scones to sell in his cafe'. If any scones are unsold at the end of the day, Duncan throws them away. The number of scones requested during a day may be modelled by a Poisson distribution with mean 27. Estimate the probability that Duncan does not have enough scones to satisfy all the requests on a particular day (*Use normal approximation to the Poisson distribution*).

Question 5 **[20 marks, 6+4+4+6]**

(a) A short-stay car park in a shopping area has spaces marked out for 90 cars. A local councillor notices that there are always some vacant spaces. He puts forward a plan to create a garden and seating area using part of the car park. This would reduce the number of parking spaces to 78.

(i) From a random sample of 14 users of the car park, 11 say that the car park will be too small if this plan is carried out. Carry out a test, at the 5% significance level, to determine whether more than half of the users of the car park think it will be too small.

(ii) The number of occupied spaces, x, in the car park is recorded on each of 16 randomly chosen occasions during shopping hours. The results may be summarised as follows:

$$\bar{x} = 59.9 \qquad s = 7.83$$

Construct a 95% confidence interval for the mean, μ, of the number of spaces occupied in the car park during shopping hours. Assume that the sample is drawn from a normal population.

(iii) The councillor claims that the value of μ is no more than 65. It is found that the number of occupied spaces during shopping hours is best modelled by a Poisson distribution with mean μ. Taking μ to be 65, use a distributional approximation to find the probability that more than 78 spaces are occupied in the car park at any one time.

(b) In a first phase of a health study in a city, a random sample of size 2000 is to be obtained. The city is comprised (broadly) of five different ethnic subpopulations that make up 40%, 30%, 10%, 10% and 10% of the city population respectively.

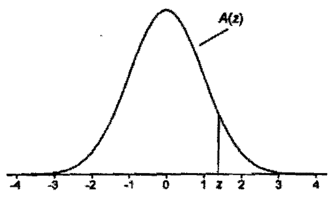
A commercial company is employed to obtain the random sample, with the instruction that the sample should reflect the ethnic composition of the city. The sample they return is summarized in the following table.

	Ethnic Subpopulation				
	1	**2**	**3**	**4**	**5**
Number in Sample	822	638	210	157	173

Using a Chi-squared test for this one-way layout, comment on whether the company have fulfilled their remit to produce a sample that reflects the ethnic composition of the city.

TABLE A.1

Cumulative Standardized Normal Distribution



$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

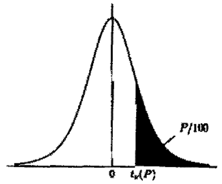
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Percentage Points of the t-Distribution

This table gives the percentage points $t_\nu(P)$ for various values of P and degrees of freedom ν, as indicated by the figure to the right.

The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geq t_\nu(P)$ is $2P/100$.

The limiting distribution of t as $\nu \to \infty$ is the normal distribution with zero mean and unit variance.



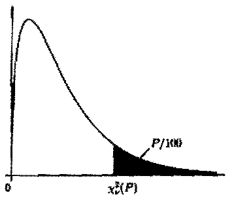
ν	Percentage points P						
	10	5	2.5	1	0.5	0.1	0.05
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Percentage Points of the χ^2-Distribution

This table gives the percentage points $\chi^2_\nu(P)$ for various values of P and degrees of freedom ν, as indicated by the figure to the right.

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geq \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.



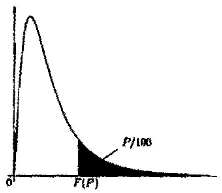
ν	Percentage points P						
	10	5	2.5	1	0.5	0.1	0.05
1	2.706	3.841	5.024	6.635	7.879	10.828	12.116
2	4.605	5.991	7.378	9.210	10.597	13.816	15.202
3	6.251	7.815	9.348	11.345	12.838	16.266	17.730
4	7.779	9.488	11.143	13.277	14.860	18.467	19.997
5	9.236	11.070	12.833	15.086	16.750	20.515	22.105
6	10.645	12.592	14.449	16.812	18.548	22.458	24.103
7	12.017	14.067	16.013	18.475	20.278	24.322	26.018
8	13.362	15.507	17.535	20.090	21.955	26.124	27.868
9	14.684	16.919	19.023	21.666	23.589	27.877	29.666
10	15.987	18.307	20.483	23.209	25.188	29.588	31.420
11	17.275	19.675	21.920	24.725	26.757	31.264	33.137
12	18.549	21.026	23.337	26.217	28.300	32.909	34.821
13	19.812	22.362	24.736	27.688	29.819	34.528	36.478
14	21.064	23.685	26.119	29.141	31.319	36.123	38.109
15	22.307	24.996	27.488	30.578	32.801	37.697	39.719
16	23.542	26.296	28.845	32.000	34.267	39.252	41.308
17	24.769	27.587	30.191	33.409	35.718	40.790	42.879
18	25.989	28.869	31.526	34.805	37.156	42.312	44.434
19	27.204	30.144	32.852	36.191	38.582	43.820	45.973
20	28.412	31.410	34.170	37.566	39.997	45.315	47.498
25	34.382	37.652	40.646	44.314	46.928	52.620	54.947
30	40.256	43.773	46.979	50.892	53.672	59.703	62.162
40	51.805	55.758	59.342	63.691	66.766	73.402	76.095
50	63.167	67.505	71.420	76.154	79.490	86.661	89.561
80	96.578	101.879	106.629	112.329	116.321	124.839	128.261

5 Percent Points of the F-Distribution

This table gives the percentage points $F_{\nu_1,\nu_2}(P)$ for $P = 0.05$ and degrees of freedom ν_1, ν_2, as indicated by the figure to the right.

The lower percentage points, that is the values $F'_{\nu_1,\nu_2}(P)$ such that the probability that $F \leq F'_{\nu_1,\nu_2}(P)$ is equal to $P/100$, may be found using the formula

$$F'_{\nu_1,\nu_2}(P) = 1/F_{\nu_1,\nu_2}(P)$$



ν_2	ν_1: 1	2	3	4	5	6	12	24	∞
2	18.513	19.000	19.164	19.247	19.296	19.330	19.413	19.454	19.496
3	10.128	9.552	9.277	9.117	9.013	8.941	8.745	8.639	8.526
4	7.709	6.944	6.591	6.388	6.256	6.163	5.912	5.774	5.628
5	6.608	5.786	5.409	5.192	5.050	4.950	4.678	4.527	4.365
6	5.987	5.143	4.757	4.534	4.387	4.284	4.000	3.841	3.669
7	5.591	4.737	4.347	4.120	3.972	3.866	3.575	3.410	3.230
8	5.318	4.459	4.066	3.838	3.687	3.581	3.284	3.115	2.928
9	5.117	4.256	3.863	3.633	3.482	3.374	3.073	2.900	2.707
10	4.965	4.103	3.708	3.478	3.326	3.217	2.913	2.737	2.538
11	4.844	3.982	3.587	3.357	3.204	3.095	2.788	2.609	2.404
12	4.747	3.885	3.490	3.259	3.106	2.996	2.687	2.505	2.296
13	4.667	3.806	3.411	3.179	3.025	2.915	2.604	2.420	2.206
14	4.600	3.739	3.344	3.112	2.958	2.848	2.534	2.349	2.131
15	4.543	3.682	3.287	3.056	2.901	2.790	2.475	2.288	2.066
16	4.494	3.634	3.239	3.007	2.852	2.741	2.425	2.235	2.010
17	4.451	3.592	3.197	2.965	2.810	2.699	2.381	2.190	1.960
18	4.414	3.555	3.160	2.928	2.773	2.661	2.342	2.150	1.917
19	4.381	3.522	3.127	2.895	2.740	2.628	2.308	2.114	1.878
20	4.351	3.493	3.098	2.866	2.711	2.599	2.278	2.082	1.843
25	4.242	3.385	2.991	2.759	2.603	2.490	2.165	1.964	1.711
30	4.171	3.316	2.922	2.690	2.534	2.421	2.092	1.887	1.622
40	4.085	3.232	2.839	2.606	2.449	2.336	2.003	1.793	1.509
50	4.034	3.183	2.790	2.557	2.400	2.286	1.952	1.737	1.438
100	3.936	3.087	2.696	2.463	2.305	2.191	1.850	1.627	1.283
∞	3.841	2.996	2.605	2.372	2.214	2.099	1.752	1.517	1.002