UNIVERSITY OF SWAZILAND

FINAL EXAMINATION PAPER

PROGRAMME	**DIPLOMA IN AGRICLTURE** **DIPLOMA IN AGRICULTURAL EDUCATION** **DIPLOMA IN HOME ECONOMICS** **DIPLOMA IN HOME ECONOMICS EDUCATION** **REMEDIAL IN AGRICULTURE**
COURSE CODE:	**AEM 201**
TITLE OF PAPER:	**ELEMENTARY STATISTICS**
TIME ALLOWED:	**TWO HOURS (2HRS)**
REQUIREMENTS:	**CALCULATOR, STATISTICAL TABLES AND GRAPH PAPER**
INSTRUCTIONS:	**ANSWER QUESTION ONE AND ANY OTHER TWO QUESTIONS.**

DO NOT OPEN THIS PAPER UNTIL PERMISSION HAS BEEN GRANTED BY THE CHIEF INVIGILATOR

Question 1

a) The data given below are the numbers (in thousands) of farms in each of 50 states, given in order of increasing numbers of farms.

1	1	2	3	3	4	4	6	7	8	8	9	9
14	14	17	21	24	24	27	28	33	36	36	37	38
39	42	46	49	50	50	52	57	57	61	70	71	73
78	79	82	87	88	93	96	99	109	115	160		

Compute the median, mean, standard deviation and Comment on the skewness (spread) of this distribution. (10 marks)

b) Two human resource managers in different citrus companies want to compare the lengths of time that employees stay in their organisations before leaving. They consulted their records on 31 March and looked at all the employees who left the organisations in the previous three months. The data are shown in the table.

Number of employees

Length of service (months)	*Company A*	*Company B*
Less than 1	2	40
1 but less than 2	3	20
2 but less than 3	7	18
3 but less than 6	20	22
6 but less than 12	33	20
12 but less than 24	20	20
24 but less than 36	10	20
36 or more	5	40

From the above table, estimate;
i) the median length of service in Company A,
ii) the median length of service in Company B,
iii) the inter-quartile range of length of service in Company A,
iv) the inter-quartile range of length of service in Company B.
v) compare the lengths of service at the two companies.

(15 Marks)

Question 2

A sample of 10 sea bass was caught by a fisheries scientist who then measured their length x (in millimetres) and their weight y (in grams). The data are given in the table below.

Length (x)	387	366	329	293	273	268	294	198	185	169
Weight (y)	720	680	480	330	270	220	380	108	89	68

a) Calculate a measure of the magnitude of the relationship between x and y variables.
b) Calculate the least-squares estimates of the parameters of the regression line.
c) Comment on the appropriateness of the regression line estimated in part (a) as a model for the relationship between the weights and lengths of sea bass.
d) Calculate the standard error of the regression line.

(5 + 6 + 4 + 5 Marks)

Question 3

a) Write down a formula for P(A|B), the conditional probability of an event A given an event B. (You may assume that the probability of B is non-zero.)

(3 Marks)

b) State what is meant by saying that two events A and B are independent
   i) in terms of P(A and B);
   ii) in terms of P(A|B).

(4 Marks)

c) A farm machine has two components X and Y and have respective probabilities 3/4 and 7/8 of functioning correctly. They function independently of one another. Two devices are constructed using such components. Device 1 works only if both X and Y function correctly, Device 2 works whenever at least one of X and Y functions correctly.
   i) Find the probability that Device 1 works.
   ii) Find the probability that Device 2 works.

(6 marks)

d) Suppose Device 1 works. Find the three probabilities
   i) that X is functioning correctly,
   ii) that only X is functioning correctly,
   iii) that both X and Y are functioning correctly.

Find the same probabilities, supposing instead that Device 2 works. (7 Marks)

Question 4

A supermarket has a policy of only buying tomatoes from growers who can supply tomatoes that have a mean diameter of 3.0 cm and a standard deviation of no more than 0.5 cm. A representative of the supermarket goes to visit a potential new supplier and selects a random sample of 16 tomatoes from the tomato grower's greenhouse. The diameter of each tomato is measured and the data are as follows, recorded for convenience in ascending order.

2.2, 2.3, 2.5, 2.6, 2.6, 2.7, 2.9, 3.0, 3.2, 3.3, 3.4, 3.6, 3.6, 3.8, 3.8, 3.9

a) By constructing suitable confidence intervals, analyse these data to establish whether the tomatoes provided by the grower will meet the supermarket's requirements, clearly stating any assumptions on which your analysis depends. Write a short report to the board of directors outlining your recommendations concerning whether or not to use this tomato grower to supply tomatoes for sale in the supermarket.

(15 marks)

b) The supermarket representative suggests that the simplest way to select the sample would be to pick two tomato plants at random and select eight tomatoes at random from each. Comment on the suitability of this method.

(5 Marks)

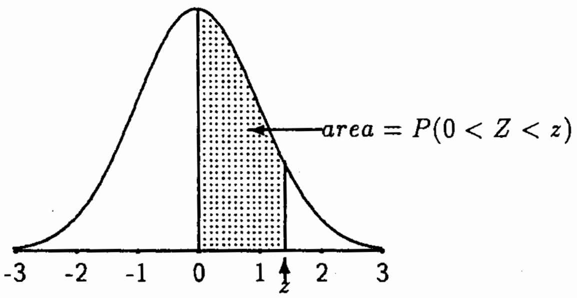
Confidence Intervals

Parameter	Assumptions	Endpoints
μ	$N(\mu, \sigma^2)$ or n large σ^2 known	$\overline{x} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$
μ	$N(\mu, \sigma^2)$ σ^2 unknown	$\overline{x} \pm t_{\alpha/2}(n-1) \dfrac{s}{\sqrt{n}}$
$\mu_X - \mu_Y$	$N(\mu_X, \sigma_X^2)$ $N(\mu_Y, \sigma_Y^2)$ σ_X^2, σ_Y^2 known	$\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}$
$\mu_X - \mu_Y$	Variances unknown, large samples	$\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\dfrac{s_x^2}{n} + \dfrac{s_y^2}{m}}$
$\mu_X - \mu_Y$	$N(\mu_X, \sigma_X^2)$ $N(\mu_Y, \sigma_Y^2)$ $\sigma_X^2 = \sigma_Y^2$, unknown	$\overline{x} - \overline{y} \pm t_{\alpha/2}(n+m-2)s_p\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}$, $s_p = \sqrt{\dfrac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}$
$\mu_D = \mu_X - \mu_Y$	X and Y normal, but dependent	$\overline{d} \pm t_{\alpha/2}(n-1)\dfrac{s_d}{\sqrt{n}}$
σ^2	$N(\mu, \sigma^2)$	$\dfrac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}$
$\dfrac{\sigma_X^2}{\sigma_Y^2}$	$N(\mu_X, \sigma_X^2)$ $N(\mu_Y, \sigma_Y^2)$	$\dfrac{s_x^2/s_y^2}{F_{\alpha/2}(n-1, m-1)}, \; F_{\alpha/2}(m-1, n-1)\dfrac{s_x^2}{s_y^2}$
p	$b(n, p)$ n is large	$\dfrac{y}{n} \pm z_{\alpha/2}\sqrt{\dfrac{(y/n)(1-y/n)}{n}}$
$p_1 - p_2$	$b(n_1, p_1)$ $b(n_2, p_2)$	$\widehat{p}_1 - \widehat{p}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \dfrac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}$, $\widehat{p}_1 = y_1/n_1, \; \widehat{p}_2 = y_2/n_2$

Tests of Hypotheses

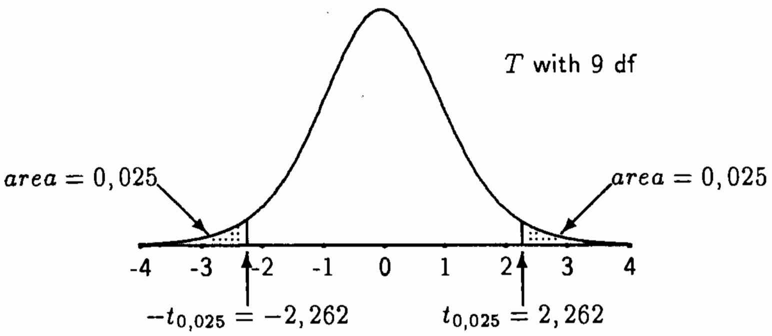
Hypotheses	Critical Region
H_0: $\mu = \mu_0$ H_1: $\mu > \mu_0$ σ^2 known	$z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} \ge z_\alpha$
H_0: $\mu = \mu_0$ H_1: $\mu > \mu_0$ σ^2 unknown	$z = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}} \ge t_\alpha(n-1)$
H_0: $\mu_X - \mu_Y = 0$ H_1: $\mu_X - \mu_Y > 0$ σ_X^2, σ_Y^2 known	$z = \dfrac{\overline{x} - \overline{y} - 0}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \ge z_\alpha$
H_0: $\mu_X - \mu_Y = 0$ H_1: $\mu_X - \mu_Y > 0$ Variances unknown, large samples	$z = \dfrac{\overline{x} - \overline{y} - 0}{\sqrt{s_x^2/n + s_y^2/m}} \ge z_\alpha$
H_0: $\mu_X - \mu_Y = 0$ H_1: $\mu_X - \mu_Y > 0$ $\sigma_X^2 = \sigma_Y^2$, unknown	$t = \dfrac{\overline{x} - \overline{y} - 0}{s_p\sqrt{1/n + 1/m}} \ge t_\alpha(n+m-2)$ $s_p = \sqrt{\dfrac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}$
H_0: $\mu_D = 0$ H_1: $\mu_D > 0$	$t = \dfrac{\overline{d} - 0}{s_d/\sqrt{n}} \ge t_\alpha(n-1)$
H_0: $\sigma^2 = \sigma_0^2$ H_1: $\sigma^2 > \sigma_0^2$	$\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2} \ge \chi^2_\alpha(n-1)$
H_0: $\sigma_X^2/\sigma_Y^2 = 1$ H_1: $\sigma_X^2/\sigma_Y^2 > 1$	$F = \dfrac{s_x^2}{s_y^2} \ge F_\alpha(n-1, m-1)$
H_0: $p = p_0$ H_1: $p > p_0$	$z = \dfrac{y/n - p_0}{\sqrt{p_0(1-p_0)/n}} \ge z_\alpha$
H_0: $p_1 - p_2 = 0$ H_1: $p_1 - p_2 > 0$	$z = \dfrac{y_1/n_1 - y_2/n_2 - 0}{\sqrt{\left(\dfrac{y_1 + y_2}{n_1 + n_2}\right)\left(1 - \dfrac{y_1 + y_2}{n_1 + n_2}\right)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \ge z_\alpha$

Table of Standard Normal Curve Areas



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	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Table of Critical Values of T



df	$t_{0,05}$	$t_{0,025}$	$t_{0,01}$	$t_{0,005}$
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.045	2.462	2.756
30	1.697	2.042	2.457	2.750

df	$t_{0,05}$	$t_{0,025}$	$t_{0,01}$	$t_{0,005}$
40	1.684	2.021	2.423	2.704
50	1.676	2.009	2.403	2.678
60	1.671	2.000	2.390	2.660
70	1.667	1.994	2.381	2.648
80	1.664	1.990	2.374	2.639
90	1.662	1.987	2.368	2.632
100	1.660	1.984	2.364	2.626
110	1.659	1.982	2.361	2.621
120	1.658	1.980	2.358	2.617
130	1.657	1.978	2.355	2.614
140	1.656	1.977	2.353	2.611
150	1.655	1.976	2.351	2.609
160	1.654	1.975	2.350	2.607
170	1.654	1.974	2.348	2.605
180	1.653	1.973	2.347	2.603
190	1.653	1.973	2.346	2.602
200	1.653	1.972	2.345	2.601
210	1.652	1.971	2.344	2.599
220	1.652	1.971	2.343	2.598
230	1.652	1.970	2.343	2.597
240	1.651	1.970	2.342	2.596
250	1.651	1.969	2.341	2.596
260	1.651	1.969	2.341	2.595
270	1.651	1.969	2.340	2.594
280	1.650	1.968	2.340	2.594
290	1.650	1.968	2.339	2.593
300	1.650	1.968	2.339	2.592
310	1.650	1.968	2.338	2.592
320	1.650	1.967	2.338	2.591
330	1.649	1.967	2.338	2.591
∞	1.645	1.96	2.33	2.58