UNIVERSITY EXAMINATIONS: 2020/2021

FINAL EXAMINATION FOR THE DEGREE OF BACHELOR OF COMMERCE

STA 2101: BUSINESS STATISTICS II

CMS 301-F: ADVANCED BUSINESS STATISTICS

OPEN BOOK: FULLTIME

DATE: AUGUST 2020 TIME: 3 HOURS

INSTRUCTIONS: Answer ***BOTH*** Questions

QUESTION ONE [25 MARKS]

(a) Last year Githunguri Dairies decided to enter the yogurt market, and it began cautiously by producing, distributing, and marketing a single flavor—a blueberry-flavored yogurt that it calls *Smooth*. The company's initial venture into the yogurt market has been very successful; sales of *Smooth* are higher than expected, and consumers' ratings of the product have a mean of 80 and a standard deviation of 25 on a 100-point scale for which 100 is the most favorable score and zero is the least favorable score. Past experience has also shown Githunguri Dairies that a consumer who rates one of its products with a score greater than 75 on this scale will consider purchasing the product, and a score of 75 or less indicates that the consumer will not consider purchasing the product.

Emboldened by the success and popularity of its blueberry-flavored yogurt, Githunguri Dairies management is now considering the introduction of a second flavor. Githunguri's marketing department is pressing to extend the product line through the introduction of a strawberry-flavored yogurt that would be called *Energy*, but senior managers are concerned about whether or not *Energy* will increase Githunguri's market share by appealing to potential customers who do not like *Smooth*. That is, the goal in offering the new product is to increase the market share rather than cannibalize existing sales of *Smooth*. The marketing department has proposed giving tastes of both *Smooth* and *Energy* to a simple random sample of 50 customers and asking each of them to rate the two yogurts on the 100-point scale. If the mean score given to *Smooth* by this sample of consumers is 75 or less, Githunguri's senior management believes the sample can be used to assess whether *Energy* will appeal to potential customers who do not like *Smooth*.

(i) Calculate the probability that the mean score of *Smooth* given by the simple random sample of Githunguri Dairies customers will be 75 or less. [2 marks]

(ii) If the Marketing Department increases the sample size to 150, what is the probability that the mean score of *Smooth* given by the simple random sample of Githunguri Dairies customers will be 75 or less? [2 marks]

(iii) Explain to Githunguri Dairies senior management why the probability that the mean score of *Smooth* given by the simple random sample of Githunguri Dairies customers will be 75 or less differs for these two sample sizes. [2 marks]

(b) Geosoft Research, a consumer research organization, conducts surveys designed to evaluate a wide variety of products and services available to consumers. In one particular study, Geosoft looked at consumer satisfaction with the performance of automobiles produced by a major Thika manufacturer. A questionnaire sent to owners of one of the manufacturer's full-sized cars revealed several complaints about early transmission problems. To learn more about the transmission failures, Geosoft used a sample of actual transmission repairs provided by a transmission repair firm in the Thika area. The following data show the actual number of kilometers driven for 50 vehicles at the time of transmission failure.

85,092	32,609	59,465	77,437	32,534	64,090	32,464	59,902
39,323	89,641	94,219	116,803	92,857	63,436	65,605	85,861
64,342	61,978	67,998	59,817	101,769	95,774	121,352	69,568
74,276	66,998	40,001	72,069	25,066	77,098	69,922	35,662
74,425	67,202	118,444	53,500	79,294	64,544	86,813	116,269
37,831	89,341	73,341	85,288	138,114	53,402	85,586	82,256
77,539	88,798						

(i) Use appropriate descriptive statistics to summarize the transmission failure data. [3 marks]

(ii) Develop a 95% confidence interval for the mean number of kilometers driven until transmission failure for the population of automobiles with transmission failure. Provide a managerial interpretation of the interval estimate. [2 marks]

(iii) Discuss the implication of your statistical findings in terms of the belief that some owners of the automobiles experienced early transmission failures. [2 marks]

(iv) How many repair records should be sampled if the research firm wants the population mean number of kilometers driven until transmission failure to be estimated with a margin of error of 5000 kilometers? Use 95% confidence. [2 marks]

(v) What other information would you like to gather to evaluate the transmission failure problem more fully? [2 marks]

(c) A sociologist was hired by a large city hospital to investigate the relationship between the numbers of unauthorized days that employees are absent per year and the distance (miles) between home and work for the employees. A sample of 10 employees was chosen, and the following data were collected.

Distance to Work (miles)	Number of Days Absent
1	8
3	5
4	8
6	7
8	6
10	3
12	5
14	2
14	4
18	2

(i) Develop a scatter diagram for these data. Does a linear relationship appear reasonable? Explain. [3 marks]

(ii) Develop the least squares estimated regression equation. [5 marks]

QUESTION TWO [25 MARKS]

(a) Data Trends Ltd., a consulting firm, advises its clients about sampling and statistical procedures that can be used to control their manufacturing processes. In one particular application, a client gave Quality associates a sample of 800 observations taken during a time in which that client's process was operating satisfactorily. The sample standard deviation for these data was 0.21; hence, with so much data, the population standard deviation was assumed to be 0.21. Quality associates then suggested that random samples of size 30 be taken periodically to monitor the process on an ongoing basis. By analyzing the new samples, the client could quickly learn whether the process was operating satisfactorily. When the process was not operating satisfactorily, corrective action could be taken to eliminate the problem. The design specification indicated the mean for the process should be 12. The hypothesis test suggested by Quality associates follows.

$$H_0 : \mu = 12$$
$$H_A : \mu \neq 12$$

Corrective action will be taken any time H_0 is rejected.

The following samples were collected at hourly intervals during the first day of operation of the new statistical process control procedure.

Sample 1	Sample 2	Sample 3	Sample 4
11.55	11.62	11.91	12.02
11.62	11.69	11.36	12.02
11.52	11.59	11.75	12.05
11.75	11.82	11.95	12.18
11.90	11.97	12.14	12.11
11.64	11.71	11.72	12.07
11.80	11.87	11.61	12.05
12.03	12.10	11.85	11.64
11.94	12.01	12.16	12.39
11.92	11.99	11.91	11.65
12.13	12.20	12.12	12.11
12.09	12.16	11.61	11.90
11.93	12.00	12.21	12.22
12.21	12.28	11.56	11.88
12.32	12.39	11.95	12.03
11.93	12.00	12.01	12.35
11.85	11.92	12.06	12.09
11.76	11.83	11.76	11.77
12.16	12.23	11.82	12.20
11.77	11.84	12.12	11.79
12.00	12.07	11.60	12.30
12.04	12.11	11.95	12.27
11.98	12.05	11.96	12.29
12.30	12.37	12.22	12.47
12.18	12.25	11.75	12.03
11.97	12.04	11.96	12.17
12.17	12.24	11.95	11.94
11.85	11.92	11.89	11.97
12.30	12.37	11.88	12.23
12.15	12.22	11.93	12.25

(i) Conduct a hypothesis test for each sample at the 1% level of significance and determine what action, if any, should be taken. Provide the test statistic and p-value for each test. [7 marks]

(ii) Compute the standard deviation for each of the four samples. Does the assumption of 0.21 for the population standard deviation appear reasonable? [5 marks]

(iii) Discuss the implications of changing the level of significance to a larger value. What mistake or error could increase if the level of significance is increased? [2 marks]

(b) The following are quality control data for a manufacturing process at ChemBio Chemical Company. The data show the temperature in degrees centigrade at five points in time during a manufacturing cycle. The company is interested in using control charts to monitor the temperature of its manufacturing process. Construct the $\bar{x}$ chart and R chart. What conclusions can be made about the quality of the process? [6 marks]

Sample	$\bar{x}$	R	Sample	$\bar{x}$	R
1	95.72	1.0	11	95.80	.6
2	95.24	.9	12	95.22	.2
3	95.18	.8	13	95.56	1.3
4	95.44	.4	14	95.22	.5
5	95.46	.5	15	95.04	.8
6	95.32	1.1	16	95.72	1.1
7	95.40	.9	17	94.82	.6
8	95.44	.3	18	95.46	.5
9	95.08	.2	19	95.60	.4
10	95.50	.6	20	95.74	.6

(c) Health insurance benefits vary by the size of the company. The sample data below show the number of companies providing health insurance for small, medium, and large companies. For purposes of this study, small companies are companies that have fewer than 100 employees. Medium-sized companies have 100 to 999 employees, and large companies have 1000 or more employees.

The questionnaire sent to 225 employees asked whether or not the employee had health insurance and then asked the employee to indicate the size of the company.

	Size of the Company		
Health Insurance	Small	Medium	Large
Yes	36	65	88
No	14	10	12

Conduct a test of independence to determine whether health insurance coverage is independent of the size of the company. What is the p-value? Using a 5% level of significance, what is your conclusion? [5 marks]

t - Table with right tail probabilities

f\p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	4.3178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
inf	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905

Area between 0 and 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

Control Chart Factors

n	d2	A2	d3	D3	D4
2	1.128	1.880	0.853	0.000	3.267
3	1.693	1.023	0.888	0.000	2.575
4	2.059	0.729	0.880	0.000	2.282
5	2.326	0.577	0.864	0.000	2.114
6	2.534	0.483	0.848	0.000	2.004
7	2.704	0.419	0.833	0.076	1.924
8	2.847	0.373	0.820	0.136	1.864
9	2.970	0.337	0.808	0.184	1.816
10	3.078	0.308	0.797	0.223	1.777
11	3.173	0.285	0.787	0.256	1.744
12	3.258	0.266	0.778	0.283	1.717
13	3.336	0.249	0.770	0.307	1.693
14	3.407	0.235	0.763	0.328	1.672
15	3.472	0.223	0.756	0.347	1.653
16	3.532	0.212	0.750	0.363	1.637
17	3.588	0.203	0.744	0.378	1.622
18	3.640	0.194	0.739	0.391	1.609
19	3.689	0.187	0.733	0.404	1.596
20	3.735	0.180	0.729	0.415	1.585
21	3.778	0.173	0.724	0.425	1.575
22	3.819	0.167	0.720	0.435	1.565
23	3.858	0.162	0.716	0.443	1.557
24	3.895	0.157	0.712	0.452	1.548
25	3.931	0.153	0.708	0.459	1.541