國立成功大學104學年度碩士班招生考試試題

系所組別：資訊管理研究所甲組

考試科目：統計學

考試日期：0212，節次：3

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

1. (5%) Tell in each of the following instances whether the study uses an independent samples or a matched pairs design.
   (a) Two computing algorithms are compared in terms of the CPU times required to do the same six test problems.
   (b) A survey is conducted of teens from inner city schools and suburban schools to compare the proportion who have tried drugs.
   (c) An advertising agency has come up with two different TV commercials for a household detergent. To determine which one is more effective, a test is conducted in which a sample of 100 adults is randomly divided into two groups. Each group is shown a different commercial, and the people in the group are asked to score the commercial.

2. (10%) Find the variance of X, the random variable with probability mass function

$$p(x)=\begin{cases}(|x-3|+1)/28 & x=-3,-2,-1,0,1,2,3\\ 0 & \text{otherwise.}\end{cases}$$

3. (10%) The distribution function for the duration of a certain soap opera is

$$F(x)=\begin{cases}1-\dfrac{16}{x^2} & \text{if } x\geq 4\\ 0 & \text{if } x<4.\end{cases}$$

   (a) Find $E(X)$.
   (b) Show that $Var(X)$ does not exist.

4. (10%) An engineer suspects that the temperature inside an oven is not as uniform as when it was new, at which time the temperature varied $\pm 10^\circ$F around its setting. (Taking the range of a normal distribution to be roughly $\pm 2\sigma$, this translates into $\sigma = 5^\circ$F.) To verify his suspicion, he takes 20 measurements in different parts of the oven. He wants a rule which decides that the true $\sigma > 5$ if the sample standard deviation of the measurements exceeds $5c$, where $c > 0$ is a suitably chosen constant. The rule must not have more than a 10% chance of making a wrong decision, i.e., deciding that $\sigma > 5$ when, in fact, $\sigma = 5$.
   (a) Find the value of c.
   (b) Based on this value of c, does the rule decide that $\sigma > 5$ if the sample standard deviation of the engineer's measurements is $s = 7.5^\circ$F?

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

5. (15%) In order to shorten the cycle time of an assembly process, an industrial engineer has simplified the operations and designed a new fixture. With the old method, the average time for assembly is 10 minutes. To test his claim an experiment is conducted in which 15 workers trained in the new method participate.

(a) Set up the hypotheses to show that the average assembly time with the new method is less than 10 minutes.

(b) Suppose that the sample mean for the 15 workers is 8.7 minutes. If $\sigma = 2$ minutes, is there statistically significant evidence that the average time is reduced? Use $\alpha = 0.05$.

(c) The industrial engineer claims that the new method will reduce the average time by at least 1.5 minutes. What chance (power) does this experiment have of detecting the claimed improvement?

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

6. An engineer would like to use two variables, temperature (x_1, °C) and catalyst feed rate (x_2, lb/h), to predict the reaction rate (Y) of a chemical process. The original data is not given here but there is a table (generated by MINITAB) given when fitting a linear regression model between Y and (x_1, x_2). Answer the following questions by reading the table.

Table 1: Regression Analysis: Y v.s. (x_1, x_2)

```
The regression equation is
Y = 99.4 + 2.82 x1 + 0.45 x2

Predictor    Coef  SE Coef     T      P
Constant    99.43    39.30  2.53  0.039
x1         2.8245   0.3804  7.42  0.000
x2          0.447    1.158  0.39  0.711

S = 7.25429   R-Sq = 91.1%   R-Sq(adj) = 88.6%

Analysis of Variance

Source          DF      SS      MS      F      P
Regression       2  3782.8  1891.4  35.94  0.000
Residual Error   7   368.4    52.6
Total            9  4151.2
```

(1) (5%) Explain how the coefficients in the fitted regression equation was obtained?

(2) (5%) Interpret the obtained fitted regression equation.

(3) (5%)The engineer also tried to find a simple linear regression line between Y and two independent variables x_1 and x_2, respectively. The outputs are given in Tables 2 and 3. He notices that the values of the coefficient of determination, R^2, in Table 1 is higher than that from Tables 2 and 3 and he wonders if this is correct? Explain why this is true?

Table 2: Regression Analysis: Y v.s. x_1

Regression Analysis: Y versus x1

```
The regression equation is
Y = 110 + 2.75 x1

Predictor    Coef  SE Coef     T      P
Constant   110.22    26.16  4.21  0.003
x1         2.7477   0.3067  8.96  0.000

S = 6.85775   R-Sq = 90.9%   R-Sq(adj) = 89.8%

Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  3775.0  3775.0  80.27  0.000
Residual Error   8   376.2    47.0
Total            9  4151.2
```

Table 3: Regression Analysis: Y v.s. x_2

Regression Analysis: Y versus x2

```
The regression equation is
Y = 382 - 4.04 x2

Predictor    Coef  SE Coef      T      P
Constant   382.23    26.94  14.19  0.000
x2         -4.041    2.751  -1.47  0.180

S = 20.2150   R-Sq = 21.2%   R-Sq(adj) = 11.4%

Analysis of Variance

Source          DF      SS     MS     F      P
Regression       1   882.0  882.0  2.16  0.180
Residual Error   8  3269.2  408.6
Total            9  4151.2
```

(4) (5%)After a preliminary analysis, the engineer is curious about if he should use both independent variables for this analysis. Determine if x_2 variable should be kept in the model if x_1 variable is already in the regression analysis?

(5) (5%)The engineer would like to use these two process variables (x_1, x_2) for predicting if a experiment is going to be successful or not (i.e., the response variable is binary and has two outcomes). However, his colleague suggests him that this is not a correct approach. When the dependent variable (Y) is binary, a linear regression analysis won't provide a satisfactory result. Do you agree with this statement? Why or Why not?

國立成功大學104學年度碩士班招生考試試題

系所組別：資訊管理研究所甲組

考試科目：統計學

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

7. In a factory, a manager is trying to figure out the effect of three assembly methods (called Method 1, 2, and 3) on the assembling completion time of a product that the factory produces. Three operators are chosen for this experiment. The manager realizes because of different levels of experience in assembling the product. It will be better to consider a Randomized Complete Block Design (RCBD).

(1) (5%) State that how a RCBD is conducted under this scenario?

(2) (5%) After the experiment was carried out, the data is then analyzed and the ANOVA table is given. Fill in the blanks in the ANOVA table.

Source	DF	SS	MS	F
Method				
Operator	2	32.20		
Error		1.36		
Total	8	53.02		

(3) (5%) Write down your null hypothesis and make conclusions. What are the assumptions regarding the obtained data for making this test?

(4) (6%) Suppose a data analyst does not aware that this data was obtained from a RCBD but analyzing it using regular one-way ANOVA (assembly method is the factor). How would that change the ANOVA table? Fill the ANOVA table again under this scenario.

Source	DF	SS	MS	F
Method				
Error				
Total	8	53.02		

(5) (4%) Following the previous question (4), test if the assembly method is still a significant factor that affects the assembling completion time? Compare your conclusion to that from question (3).

編號： 266

國立成功大學104學年度碩士班招生考試試題

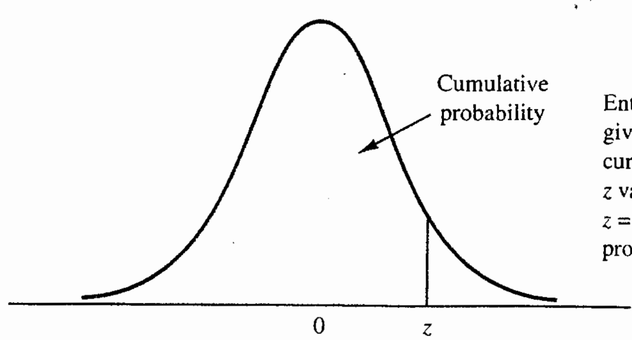
系所組別：資訊管理研究所甲組

考試科目：統計學

考試日期：0212，節次：3

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

CUMULATIVE PROBABILITIES FOR THE STANDARD NORMAL DISTRIBUTION

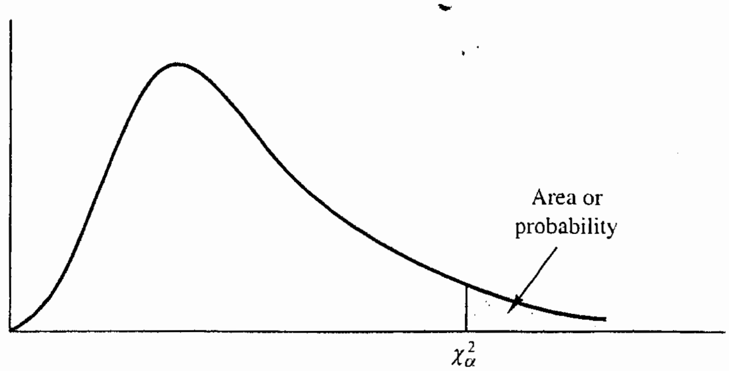


Entries in the table give the area under the curve to the left of the z value. For example, for $z = 1.25$, the cumulative probability is .8944.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9913
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9986	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

CHI-SQUARE DISTRIBUTION



Entries in the table give χ^2_α values, where α is the area or probability in the upper tail of the chi-square distribution. For example, with 10 degrees of freedom and a .01 area in the upper tail, $\chi^2_{.01} = 23.209$.

Degrees of Freedom	Area in Upper Tail									
	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	.000	.000	.001	.004	.016	2.706	3.841	5.024	6.635	7.879
2	.010	.020	.051	.103	.211	4.605	5.991	7.378	9.210	10.597
3	.072	.115	.216	.352	.584	6.251	7.815	9.348	11.345	12.838
4	.207	.297	.484	.711	1.064	7.779	9.488	11.143	13.277	14.860
5	.412	.554	.831	1.145	1.610	9.236	11.070	12.832	15.086	16.750
6	.676	.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.335

國立成功大學104學年度碩士班招生考試試題

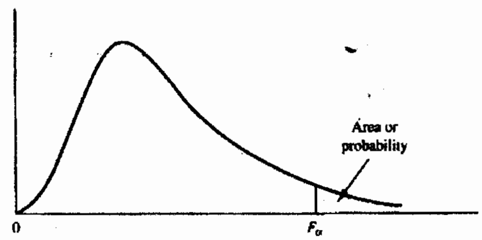
系所組別：資訊管理研究所甲組

考試科目：統計學

考試日期：0212，節次：3

※ 考生請注意：本試題可使用計算機。 請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

F-distribution table



Denominator Degrees of Freedom	Area in Upper Tail	Numerator Degrees of Freedom											
		1	2	3	4	5	6	7	8	9	10	15	20
1	.10	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74
	.05	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	245.95	248.02
	.025	647.79	799.48	864.15	899.60	921.83	937.11	948.20	956.64	963.28	968.63	984.87	993.08
	.01	4052.18	4999.34	5403.53	5624.26	5763.96	5858.95	5928.33	5980.95	6022.40	6055.93	6156.97	6208.66
2	.10	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44
	.05	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.43	39.45
	.01	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.43	99.45
3	.10	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18
	.05	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.25	14.17
	.01	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	26.87	26.69
4	.10	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84
	.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56
	.01	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.20	14.02
5	.10	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.324	3.21
	.05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33
	.01	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.72	9.55
6	.10	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84
	.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17
	.01	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.56	7.40
7	.10	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59
	.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47
	.01	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.31	6.16
8	.10	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42
	.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00
	.01	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.52	5.36
9	.10	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30
	.05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67
	.01	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	4.96	4.81
10	.10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20
	.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.77
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42
	.01	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.56	4.41
11	.10	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.17	2.12
	.05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.72	2.65
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.33	3.23
	.01	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.25	4.10
12	.10	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06
	.05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.62	2.54
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.18	3.07
	.01	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.01	3.86