1. (20%) When a test is conducted to determine whether or not someone is infected with a particular virus, an incorrect test result can occur in two ways: an infected person may test negative, or a noninfected person may test positive. The latter is called a *false positive* test. It has been pointed out that the social consequences of false positive tests for the AIDS virus are particularly serious. Such a false positive test will unnecessarily "stigmatize and frighten many healthy people," because most people consider a positive AIDS test to be a sentence to ghastly suffering and death. In order to focus on the false positive rate of the test, assume throughout this question that we are dealing with a test that properly identifies all persons who really are infected with AIDS virus.

a. Assume that 5% of a population to be tested for the AIDS virus really is infected and that the test has a false positive rate of 0.5%. Find the probability that a person who tests positive really is infected.

b. Would your answer to part (a) be higher or lower if more than 5% of the population to be tested were actually infected?

c. Assume now that a low-risk population is to be tested. Specifically, assume that 0.01% of the population to be tested is actually infected with the AIDS virus and that the test has a false positive rate of 0.005%. Find the probability that a person who tests positive really is infected.

d. What would your answer to part (c) be if you assumed a false positive rate of 0.5%?

e. Summarize the implications of your findings in parts (a) through (d).

2. (20%) Let X_i, $i = 1, 2, 3$, be independent random variable associated with a Bernoulli trial by defining it as follows: $p(X_i = 1) = p(X_i = 0) = 0.5$.

a. Find the sampling distributions of the mean, median and mode of (X_1, X_2, X_3).

b. Find the expected values and variances of the statistics in part (a).

3. (20%) The management of a firm is contemplating modifying one of its products. Before making a decision, management representatives want to conduct a market survey that will enable them to estimate the proportion of potential customers who would buy the new product. They wish to estimate this proportion to within 3%, with 99% confidence. How large a sample should be drawn in each of the following circumstances?

a. Management has no idea about the value of p.

b. It is known that the proportion will be less than 20%.

c. It is known that the proportion will be somewhere between 40% and 70%.

4. (20%) The owners of two downtown restaurants (whose customers are mostly office workers on coffee breaks) each claim to serve more coffee than the other. They decide to test their claims by counting the number of cups of coffee sold for one working week. The data are presented below. (After some analysis, it is determined that the number of cups of coffee is normally distributed.)

	NUMBERS OF CUPS OF COFFEE SOLD	
WEEKDAY	RESTAURANT 1	RESTAURANT 2
MONDAY	670	640
TUESDAY	440	420
WEDNESDAY	515	500
THURSDAY	690	650
FRIDAY	825	800

Answer the following questions by referring to the appropriate formula without performing any calculations.

a. How to estimate the average difference in coffee sales, with 99% confidence.

b. How can we conclude that there is a difference between the average coffee sales of the two restaurants? (Use $\alpha = 0.01$.)

c. Is there any connection between your answers to parts (a) and (b)? Explain.

5. (20%) a. In marketing children's products, it's extremely important to produce television commercials that hold the attention of the children who view them. A psychologist hired by a marketing research firm wants to determine whether differences in attention span exist among advertisements for different types of products. Fifteen children under 10 years of age are asked to watch one 60-second commercial for one of the three types of products, and their attention spans are measured in seconds. The results are shown in the accompanying table. Do these data provide enough evidence to conclude that there are differences in attention span among the three products advertised? (Use $\alpha = 0.05$.)

TYPE OF PRODUCT ADVERTISED		
Toy/Games	Food/Candy	Children's clothing
42	55	40
45	58	55
58	52	42
53	60	52
50	57	48

SST = 220.13, SSE = 361.60

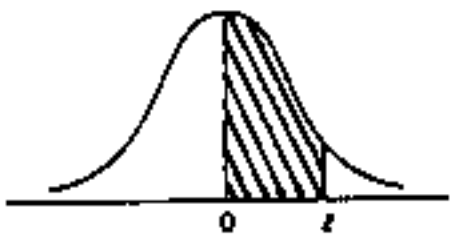
b. Upon reconsidering the experiment, the psychologist decides to redo the experiment in the following way. Three 10-year-olds, three 8-year-olds, three 6-year-olds, and three 4-year-olds are randomly assigned to watch one of the commercials, and their attention spans are measured. Do the results that follow (given in seconds) indicate, at the 5% significance level, that there are differences in the ability of the products advertised to hold children's attention?

	TYPE OF PRODUCT ADVERTISED		
Age	Toy/Games	Food/Candy	Children's clothing
10	52	60	35
8	48	58	36
6	49	54	32
4	43	52	33

Type of product sum of squares = 992.00
Age sum of squares = 68.67
Total sum of squares = 1,084.00

c. Can you explain why the psychologist decides to redo the experiment? What are the advantages and disadvantages of the two experiments?

附表



NORMAL CURVE AREAS

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

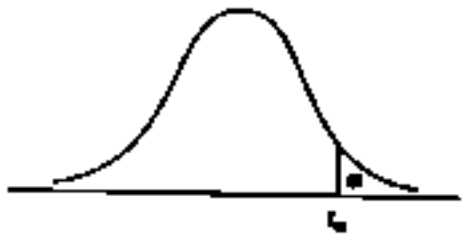
Source: Abridged from Table I of A. Hald, *Statistical Tables and Formulas* (New York: John Wiley & Sons, Inc., 1952). Reproduced by permission of A. Hald and the publisher, John Wiley & Sons, Inc.

科目：412急用統計學（資訊管理研究所乙組）

※作答前，請先核對試題、答案卷（試卷）與准考證上之所組別與考試科目是否相符！！

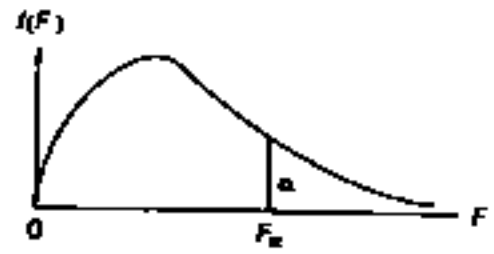
CRITICAL VALUES OF t

DEGREES OF FREEDOM	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
35	1.306	1.690	2.030	2.438	2.724
40	1.303	1.684	2.021	2.423	2.705
45	1.301	1.679	2.014	2.412	2.690
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
70	1.294	1.667	1.994	2.381	2.648
80	1.292	1.664	1.990	2.374	2.639
90	1.291	1.662	1.987	2.369	2.632
100	1.290	1.660	1.984	2.364	2.626
120	1.289	1.658	1.980	2.358	2.617
140	1.288	1.656	1.977	2.353	2.611
160	1.287	1.654	1.975	2.350	2.607
180	1.286	1.653	1.973	2.347	2.603
200	1.286	1.653	1.972	2.345	2.601
∞	1.282	1.645	1.960	2.326	2.576



SOURCE: From M. Merrington, "Table of Percentage Points of the t-Distribution," *Biometrika* 32 (1941): 300. Reproduced by permission of the Biometrika Trustees.

Percentage Points of the F Distribution, $\alpha = .05$



ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM 1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

DENOMINATOR DEGREES OF FREEDOM

SOURCE: From M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (F)-Distribution," *Biometrika* 33 (1943): 73–88. Reproduced by permission of the Biometrika Trustees.