Multiple Choice Questions (24%, 4% per question)
Please select one and only one most appropriate answer.

1. If 35% of the students in my statistics class use the Google mail (or gmail), I conclude from this that 35% of all students at the University have gmail accounts. The most important logical weakness of this conclusion would be:

A) relying on any sample instead of surveying every student.
B) using a sample that may not be representative of all students.
C) failing to correct for unconscious interviewer bias.
D) assuming cause and effect where none exists.

2. The Dean of Students mailed a survey to a total of 400 students. The sample included 100 students randomly selected from each of the freshman, sophomore, junior, and senior classes on campus last term. What sampling method was used?

A) simple random sample　　B) systematic sample
C) stratified sample　　D) cluster sample

3. A drug company wanted to test a new indigestion medication. The researchers found 400 adults aged 25-35 and randomly assigned them to two groups. The first group received the new drug, while the second received a placebo. After one month of treatment, the percentage of each group whose indigestion symptoms decreased was recorded and compared. How many levels does the treatment in this experiment have?

A) 2 (medication or placebo)　　B) 400 (number of respondents)
C) 1 (months of treatment)　　D) 10 (age span of respondents)

4. A regression intercept represents:

A) the slope of the line.　　B) the amount of change in Y given a one unit change in X.
C) the value of Y when X is equal to zero.　　D) the strength of the X-Y relationship.

5. Which is ***not*** a likely area of application of statistics in business?

A. Auditing supplier invoices for correct payment.
B. Questioning the executives' strategic decisions.
C. Looking for patterns in a large marketing database.
D. Making forecasts of several key product lines.

6. How many degrees of freedom are there in a 5x7 contingency table when the chi-square test of association is used?

A) 12　　B) 24　　C) 30　　D) 35　　E) None of the above.

Computational Questions (76%)

Please clearly show your calculations in detail.

C1. To test if a certain training course can improve performance of production line workers, the researcher randomly selected 6 workers and wrote down their performance indicators before and after the training course, as below: (The performance indicators are normally distributed)

Employee	1	2	3	4	5	6
Perf. indicator before training (X)	110	125	100	90	110	130
Perf. indicator after training (Y)	113	121	103	92	112	125

Given the sum of $x^2 = 74825$, the sum of $y^2 = 74652$, the sum of $xy = 74705$, $\alpha = 0.05$

(A) Please conduct hypothesis testing on the effectiveness of the training course using t statistics. (12%)

(B) Please calculate and present the linear regression equation for predicting Y from X. (12%)

(C) Please use ANOVA to test if the regression equation slope equals zero. (12%)

C2. 已知在總人口中帶有某遺傳基因 K 的比率為 5%，而目前測試遺傳基因 K 的檢驗對於帶有基因 K 者有 5%的機率呈現陰性，對於未帶基因 K 者有 2%機率呈現陽性。今隨機檢驗一人是否帶有基因 K，請問：

(A) 此人呈現陽性反應的機率為何？(10%)

(B) 若已知此人檢驗結果呈現陽性反應，但此人並未帶有基因 K 的機率為何？(10%)

C3. 某企業抽出 50 名員工並調查其年齡，發現平均年齡為 36 歲，標準差為 12 歲。請問：

(A) 員工平均年齡之 95%信賴區間為何？(10%)

(B) 若希望員工平均年齡的 95%信賴區間能調整到(33, 39)，請問應抽取多少樣本？(10%)

t distribution critical values

	Upper-tail probability p								
df	.25	.20	.15	.10	.05	.025	.02	.01	.005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576
	50%	60%	70%	80%	90%	95%	96%	98%	99%
	Confidence level C								

F critical values

		Degrees of freedom in the numerator						
	p	1	2	3	4	5	6	7
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91
	.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77
	.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22
	.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4
	.001	405284	500000	540379	562500	576405	585937	592873
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36
	.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67
	.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98
	.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46
	.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26
	.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99
	.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02

Degrees of freedom in the denominator