考 試 科 目	統計方法	系 所 別	統計學系	考 試 時 間	2 月 6 日(二)第 四 節

1. Please answer the following short questions.
   (a) Suppose a man roll a fair 6-sided die repeatedly until he sees the number 2 appear a total of five times, at which point he stops rolling. Let X be the total number of times he sees a number other than 2. What is the expectation of X? (3%)
   (b) Cars pass an intersection on the road at a Poisson rate of one per minute. Suppose 5% of the cars on the road are vans. Given that 10 vans have passed by in an hour, what is the expected number of total cars to have passed by in that time? (3%)
   (c) Genotypes (AA, Aa, aa) occur with probabilities (0.3, 0.4, 0.3). For $n = 3$ independent observations, the observed frequencies are (n_1, n_2, n_3). What probability distribution does n_2 alone have? (3%)
   (d) Suppose we take a random sample of 400 persons and classify each one according to their gender and color preference (red, blue, green, yellow). What hypothesis shall we want to test if we are going to perform a chi-square analysis of the resulting table? (3%)
2. The chart below records the actual high and low temperatures (unit: °F) for a selection of days in February from the weather report for a city.

Maximum	85	76	63	70	68	63
Minimum	70	68	60	65	65	61

   Assuming that the variables are normally distributed, is there sufficient evidence to conclude that there is more than a 3°F difference between average highs and lows at the 0.10 level of significance? Using p-value approach. (8%)
3. A medical research worker wants to study how male and female rats react the injection of a certain toxic substance. Suppose that 50 of 100 male rats and 160 of 400 female rats reacted strongly to the injection.
   (a) Calculate the test statistics for testing the difference between the true percentages of male and female rats which react strongly to the injection. (4%)
   (b) If analyzing the data like a 2×2 table, what test would you use for testing the difference? Also, state the hypotheses. (4%)
   (c) For the test in part (b), can you calculate the associated test statistics based on the result of part (a)? If yes, calculate the value of the test statistics; otherwise, explain your answer. (4%)
4. A national council wishes to estimate the unemployment rate (π) in Taipei city. A random sample of n persons is selected from the labor force. Let X be the number of unemployed persons in the sample.
   (a) How large a sample is necessary if the council wishes to be 95% certain that the error of estimation is within 3 percentage point when the true unemployment rate is between 0.2 and 0.4? (4%)
   (b) If n=100 and $\pi = 0.01$, can you approximate the probability that at least one unemployed persons are in the sample? If yes, calculate the value; otherwise, explain your answer. (4%)
   (c) Consider the hypothesis testing with $H_0{:}\,\pi = 1/3$ versus $H_a{:}\,\pi = 1/2$. Suppose that the rejection region for this test is $X \geq k$. If $n = 5$ and $k = 4$, calculate the probability of Type II error. (4%)

考試科目	統計方法	系所別	統計學系	考試時間	2月6日(二)第四節

(d) Refer to part (c). Using the Central Limit Theorem, find n and k such that the probability of Type I error is 0.05 and the probability of Type II error is 0.1. (8%)

5. The following table is the distribution of the number of bears spotted on 60 sightseeing tours in a national park. (Hint: $e^{-0.5} = 0.607, e^{-0.65} = 0.522, e^{-0.75} = 0.472$)

Number of tours	0	1	2	3
Number of bears	32	15	9	4

Test whether the data come from a Poisson distribution using a 1% significance level. (8%)

6. Suppose that a scientist has a bivariate dataset with $n = 5$ subjects, and suppose the sample variance of $y_1, y_2, \dots, y_n$ is 2.5, the sample variance of $x_1, x_2, \dots, x_n$ is 2.5, $\bar{x} = 3$, and the sample correlation coefficient is 0.9. A simple linear regression $Y = \beta_0 + \beta_1 X + \varepsilon$ with $\varepsilon \sim N(0, \sigma^2)$ is fit using computer software. However unfortunately a large portion of the computer outputs is missing and the scientist only has the following estimate left: $\sum_{i=1}^{5} \hat{y}_i = 25$.

(a) Fill in the entries (1) – (6) for the following table of coefficients. (6%)

Coefficient	Estimate	Std. Error	t value
Intercept	(1)	(2)	(3)
X	(4)	(5)	(6)

(b) Test whether there is a linear association between the variables using F test with $\alpha = 0.05$. (8%)

(c) Construct a 95% confidence interval for the slope. Based on this interval, would the scientist reject that $\beta_1 = 1$ at the 5% significance level? (8%)

(d) Suppose the scientist compute three confidence intervals for an individual value of Y when $X = 1$, $X = 2$, and $X = 4$. Which of the three confidence intervals gives the most precise prediction? Explain your answer. (4%)

7. An experiment is conducted and the resulting data is provided.

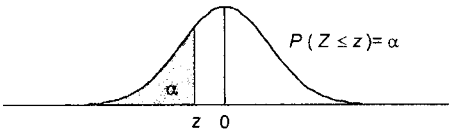
Treatment			
1	2	3	4
6	5	4	4
8	5	5	6
7	6	5	6

(a) Assuming that the data are generated from a randomized block experiment, fill in the entries (1) – (10) in the partially completed ANOVA table given below. (10%)

Source	df	SS	MS	F
Treatments	(1)	(5)	(6)	(9)
Blocks	(2)	4.2	(7)	(10)
Error	(3)	1.8	(8)	
Total	(4)	14.9		

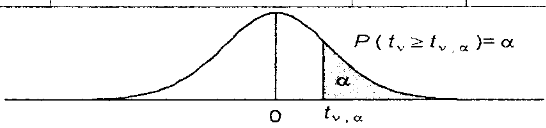
(b) Assuming now that the experimental data are generated from a one-way (independent samples) experiment, calculate the F test statistics to determine whether the treatment means differ at the 5% significance level. (4%)

考 試 科 目	統計方法	系 所 別	統計學系	考 試 時 間	2 月 6 日 (二) 第 四 節

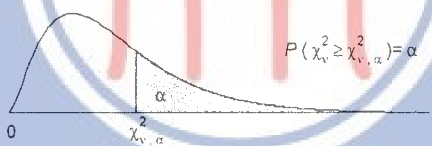


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

考試科目	統計方法	系所別	統計學系	考試時間	2 月 6 日(二) 第四節



df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
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	2.312	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



ν	Upper Tail Probability α									
	$\chi^2_{0.995}$	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$	$\chi^2_{0.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.071	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.601	9.236	11.070	12.832	15.086	16.750

備註：
一、作答於試題上者，不予計分。
二、試題請隨卷繳交。