國立成功大學 108 學年度碩士班招生考試試題

系　　所：交通管理科學系

考試科目：統計學

考試日期：0224，節次：2

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

- Unless otherwise stated, use a significance level of 5% and a confidence level of 95%
- Show all work. You may receive partial credit for partially completed problems.

1. Consider a routine screening test for a disease. Suppose the frequency of the disease in the population (base rate) is 0.5%. The test is highly accurate with a 5% false positive rate and a 10% false negative rate.

(a - 5 points) Draw the Veen diagram of the events in this problem.

(b - 5 points) Draw the tree diagram of the probabilities in this problem.

(c - 5 points) You take the test and it comes back positive. What is the probability that you have the disease?

(d - 5 points) What theorem do you use to calculate (c)? Briefly describe the theorem.

2. The data for this problem consists of pulse rate measurements on 68 males assigned at random to 6 different work tasks (Task). Here is a summary of the data in each group:

Task	n	mean	sd
1	13	31.923	4.958
2	12	31.083	5.664
3	10	35.800	5.308
4	10	38.000	6.600
5	12	29.500	6.008
6	11	28.818	4.750

(a - 3 points) Do the standard deviations satisfy the rule of thumb for safe use of ANOVA?

(b - 3 points) Is this a balanced design (yes or no)?

(c - 5 points) How many dummy variables would be required to fit the ANOVA as a multiple regression? Write out the dummy variables using an effect coding scheme.

(d - 5 points) Fill in the five pieces of missing information in the following ANOVA table.

Source	Df	Sum Sq	Mean Sq	F value
Task	_____	694.44	_____	_____
Residuals	_____	1916.08	_____	

(e - 4 points) Label the means for the different tasks $\mu 1, \mu 2, \ldots, \mu 6$. State the null and alternative hypotheses for the ANOVA.

(f - 5 points) Report the conclusions of the test at the 5% level. (You will not find the necessary denominator d.f. in the table attached, so use the nearest d.f. that is less than your denominator d.f.)

3. The economic structure of Major League Baseball (MLB) allows some teams to make substantially more money than others, which in turn allows some teams to spend much more on player salaries. These teams might therefore be expected to have better players and win more games on the field as a result. Over the course of the last 4 years (2012-2015), each of the 30 MLB teams were measured each year and the following data were collected for these 120 observations.

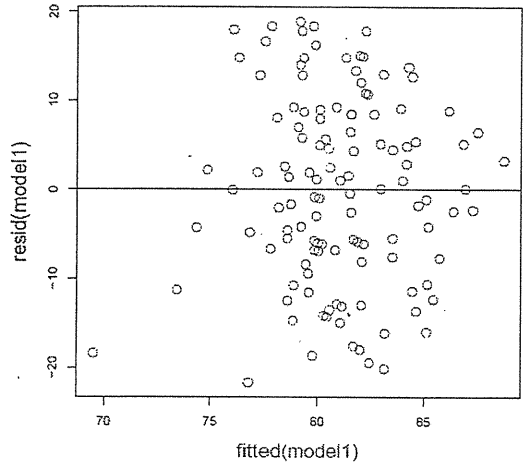
wins: number of games the team won for a specific year
payroll: opening day payroll, in millions of dollars, for the team for a specific year
AL: a binary variable for whether the team is in the American League (AL)
(14 of the 30 teams are in the AL...the rest are in the National League (NL))
year: the year in which the measurement was t

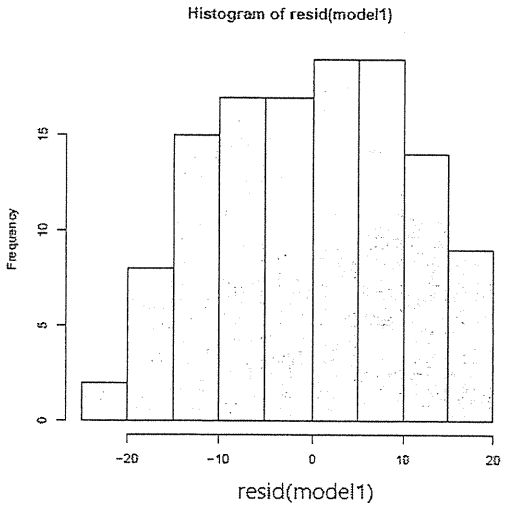
A regression model, Model 1, was run to predict the number of wins a team had from the ln(payroll), and the results in R are shown below, along with some summary statistics:
Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	44.076	11.654	3.782	0.000246
log(payroll)	7.963	2.505	3.179	0.001887

Residual standard error: 10.5649 on 118 degrees of freedom
Multiple R-squared: 0.07889, Adjusted R-squared: 0.07109
F-statistic: 10.11 on 1 and 118 DF, p-value: 0.001887
Mean of (log(mlb$payroll)) = 4.6371
SD of (log(mlb$payroll)) = 0.386

(a - 3 points) What is the estimated correlation between wins and log(payroll)?
(b - 5 points) Interpret the slope coefficient in this model. Be specific.
(c - 5 points) Briefly describe the general meaning of a p-value.
(d - 5 points) Calculate $\bar{Y}$, the mean number of wins, and s_Y , the standard deviation for the number of wins among the 120 observations.
(e - 5 points) The Red Sox are projected to spend about $180 million dollars on payroll in 2016 (next year). Provide a 95% interval for the number of games they will win next year. (t value is 1.984)
(f - 8 points) Below are the residual graphs for this model. List out all of the assumptions for this regression model, and comment on whether they are reasonable. Be specific.
(g - 4 points) Propose a different transformation on payroll to be used as a predictor. Provide 1-2 sentences of justification for why this may be a better choice for a transformation.





4. The followings are multiple choice problems. Please select the **most appropriate** answer and write down your analytical steps to get full points if necessary.

(1 - 4 points) Based on responses of 1467 respondents, a 95% confidence interval for the mean number of traffic accidents equals (6.8, 8.0). Which of the following interpretations is correct?

(a)There will be 95% probability that μ is between 6.8 and 8.0.

(b) Ninety-five percent of the values of X = number of accidents (for this sample) are between 6.8 and 8.0.

(c) If random samples of size 1467 were repeatedly selected, then 95% of the time $\bar{X}$ would be between 6.8 and 8.0.

(d) We can be 95% confident that μ is between 6.8 and 8.0.

(2 - 4 points) Which of the following statement is Correct?

(a) A continuous random variable can assume any value in a single point.

(b) The probability density function of a real-valued random variable X is the probability that X will take a value less than or equal to x.

(c) A random variable is uniformly distributed whenever the probability is proportional to the interval's length.

(d) The area under the graph of CDF and probability are identical.

(3 - 4 points) Which one of the following statements about the sample correlation coefficient, r, between two variables X and Y is false?

(a) A value of r close to 1 implies a causal relationship exists between X and Y.

(b) A value of r = 0 does not necessarily mean that X and Y are unrelated.

(c) A value of r = 0 indicates that no linear relationship exists between X and Y.

(d) A value of r = 1 indicates that a perfect positive linear relationship exists between X and Y.

(4 - 4points) Suppose we wish to test H_0 : $\mu \geq 21$ vs H_a: $\mu \leq 21$. Which of the following possible sample results gives the most evidence to support H_a

(a) $\bar{x}= 23$, $s = 3$

(b) $\bar{x}= 19$, $s = 4$

(c) $\bar{x}= 17$, $s = 7$

(d) $\bar{x}= 18$, $s = 6$

(5 - 4 points) A chi-square test involves a set of counts called "expected counts." What are the expected counts?

(a) Hypothetical counts that would occur of the alternative hypothesis were true.

(b) Hypothetical counts that would occur if the null hypothesis were true.

(c) The actual counts that did occur in the observed data.

(d) The long-run counts that would be expected if the observed counts are representative.

編號： 263

國立成功大學 108 學年度碩士班招生考試試題

系　　所：交通管理科學系

考試科目：統計學

考試日期：0224，節次：2

Upper percentage points of the F distribution with ν_1, ν_2 df

5% points

ν_1=df for the numerator

ν_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882	243.906	245.950	248.013	249.052	250.095	251.143	252.196	253.253	254.314
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5493	8.5265
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0659
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1273	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7331
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6377
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0033