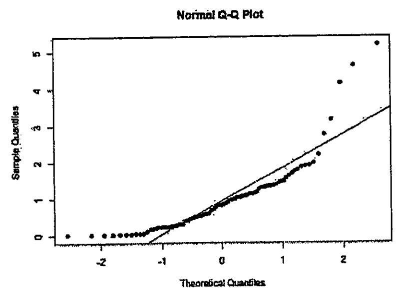
※ 考生請注意：本試題不可使用計算機。請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

一. Multiple Choice (45%)

1. A farmer plants 9 plots with a new variety of sweet potato. The average yield for these plots is 80 pounds per acre. Assume that the yield per acre for the new variety of sweet potato follows a normal distribution with unknown mean μ and standard derivation $\sigma = 9$. A 90% confidence interval for μ is:
   (a) 80 ± 5.58
   (b) 80 ± 5.88
   (c) 80 ± 4.93
   (d) 80 ± 6.98
2. A study is considered to test the hypothesis that $\mathrm{H_0}: \mu = 10$ versus $\mathrm{H_1}: \mu \neq 10$ where μ is the population mean. From a sample of 10 subjects, the test statistics is calculated to be $t = 3.25$. The p-value would be
   (a) 0.005
   (b) 0.010
   (c) 0.995
   (d) 0.990
3. The t-distribution is used to calculate a confidence interval for the population mean. If we decrease the sample size from 20 to 10, the interval would become wider because of
   (a) The change in degrees of freedom
   (b) The change in standard error.
   (c) Both (a) and (b)
   (d) None of the above
4. The histogram for hours of study per week based on a sample of 100 NCKU students is symmetric and centered at 30 hours. Which of the following statements is true?
   (a) The sample median is 30 hours.
   (b) The sample mean is 30 hours.
   (c) The skewness for the sample is zero.
   (d) All of the above.
5. Based on the normal quantile plot below, the variable can be described as:
   (a) Symmetric
   (b) Right-skewed
   (c) Left-skewed
   (d) None of the above



6. Consider the following multiple regression equation

$$\mathrm{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

where Y is the time spend on studying for the final exam in statistics, X_1 is a binary variable which takes on the value 1 for athletes and is 0 for non-athletes, and X_2 is a binary variable which takes on the value 1 for non-athletes and is 0 for athletes. If athletes spend more time studying for the final exam, then we would expect

(a) the coefficient for X_1 to have a positive sign, and for X_2 a negative sign.
(b) Both coefficients to be the same distance from the constant, one above and the other below.
(c) β_1 estimates the mean of Y for athletes, β_2 estimates the mean of Y for non-athletes.
(d) None of the above

7. Suppose the coefficient of determination R^2 for a simple linear regression is equal to 0. Which of the following statements is true?
(a) The slope coefficient is equal to 0.
(b) The correlations between the dependent and independent variables is equal to 0.
(c) The error (residual) sum of squares is equal to the total sum of squares.
(d) All of the above

8. A manufacturer claims that a particular automobile model will get 50 miles per gallon on the highway. The researchers at a consumer-oriented magazine believe that this claim is high and plan a test with a simple random sample of 36 cars. Assuming the standard deviation is 2.52 miles per gallon, what should the researchers conclude if the sample mean is 49 miles per gallon?
(a) There is not sufficient evidence to reject the manufacturer's claim; 49 miles per gallon is too close to the claimed 50 miles per gallon
(b) The manufacturer's claim should not be rejected because the p-value of .0087 is too small.
(c) The manufacturer's claim should be rejected because the sample mean is less than the claimed mean.
(d) The p-value of .0087 is sufficient evidence to reject the manufacturer's claim.

9. Which of the following are true statements?
(a) If a hypothesis test is conducted at the 1% level, there is a 1% chance of rejecting the null hypothesis.
(b) If there is sufficient evidence to reject a null hypothesis at the 10% level, then there is sufficient evidence to reject it at the 5% level.
(c) Whether to use a one- or two-tailed test is typically decided after the data are gathered.
(d) None are true.

10. Consider the null and alternative hypotheses $H_0: \hat{p} = 0.95$ versus $H_1: \hat{p} \neq 0.95$ where $\hat{p}$ is the sample proportion. These hypotheses
(a) indicate a two-tailed test.
(b) indicate a one-tailed test with a rejection area in the right tail.
(c) indicate a one-tailed test with a rejection area in the left tail.
(d) are established incorrectly.

11. A pet food manufacturer runs an experiment to determine whether three brands of dog food are equally preferred by dogs. In the experiment, 90 dogs are individually presented with three dishes of food, each containing a different brand, and their choices are noted. Tabulations show that 36 dogs go to brand A, 24 to brand B, and 30 to brand C. Which of the following statements is true?
    (a) Chi-square test for independence with test statistics $\chi^2 = 2.40$; there is not sufficient evidence at the 10% significance level.
    (b) Chi-square test for goodness of fit with test statisics $\chi^2 = 2.40$, there is not sufficient evidence at the 10% significance level.
    (c) Chi-square test for independence with test statistics $\chi^2 = 1.21$, there is not sufficient evidence at the 10% significance level.
    (d) Chi-square test for goodness of fit with test statistics $\chi^2 = 1.21$, there is not sufficient evidence at the 10% significance level.

12. Suppose we have the back-to-back stemplot below for the two random variables. Which of the following are true statements?
    I. Two random variables have the same mean.
    II. Two random variables have the same range.
    III. Two random variables have the same standard deviation.
    (a) II and III
    (b) I and II
    (c) I and III
    (d) II only

	0	348
	1	01256
843	2	29
65210	3	2557
92	4	
7552	5	6
	6	1458
6	7	09
8541	8	
90	9	

13. An Archaeopteryx is an extinct animal having feathers like a bird but also having very prominent teeth. Five fossilized specimen of this creature have been found with the lengths of the femur (a leg bone) plotted against the lengths of the humerus (a bone in the upper arm). Archeologists have long felt that there should be a positive linear correlation between the lengths of the femur and the lengths of the humerus in these extinct creatures. Using the statistical software, the following results are obtained by performing analysis for two-tailed and one-tailed test, respectively.

```
Pearson's product-moment correlation
t = 6.388, df = 3, p-value = 0.0078
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:    0.558, 0.998
sample estimates : 0.965
```

```
Pearson's product-moment correlation
t = 6.388, df = 3, p-value = 0.996
alternative hypothesis: true correlation is less than 0
95 percent confidence interval: -1.000, 0.997
sample estimates:cor 0.965
```

```
Pearson's product-moment correlation
t = 6.388, df = 3, p-value = 0.0038
alternative hypothesis: true correlation is greater than 0
95 percent confidence interval: 0.693, 1.000
sample estimates: 0.965
```

Which of the following statements is true?

(a) A positive correlation is likely because lengths of bones in any animal are always positively correlated.
(b) A positive correlation is likely because the above test yields a p-value of .0039.
(c) A positive correlation is likely because the above test yields a p-value of .0078.
(d) There is no significant positive correlation at the 1% level of significance as the P-value is .039.

14. To survey the opinions of the students in a class, a teacher plans to select every twenty-fifth student entering the classroom in the morning. Assuming there are no absences, will this sampling plan result a simple random sample of students attending the classroom?
(a) Yes, because every students has the same chance of being selected.
(b) Yes, but only if there is a single entrance to the classroom.
(c) No, because not every sample of the intended size has an equal chance of being selected.
(d) Yes, because this is an example of systematic sampling, which is a special case of simple random sampling.

15. The death rate from a particular form of cancer is 50% during the first year. When treated with an experimental drug, only 21 out of 49 patients die during the initial year. Is this strong evidence to claim that the new medication reduces the mortality rate?
(a) Yes, because the p-value is 0.3174
(b) Yes, because the p-value is 0.1587.
(c) No, because the p-value is only 0.3174.
(d) No, because the p-value is above 0.1000.

二. Problems

1. A study is done to determine the effects of removing a renal blockage in patients whose renal function is impaired because of advanced metatstatic malignancy of nonurologic cause. The arterial blood pressure of a random sample of 9 patients is measured before and after surgery for treatment of the blockage yielded the following data:

Before	150	132	130	116	107	100	101	96	78
After	90	102	80	82	90	94	84	93	80

(a) Please state the hypothesis and all the assumptions needed. (4%)
(b) Please conduct the test and state your conclusion. (6%)

2. Please answer the following problems.
I. A researcher hypothesizes that final exam scores were different across the 5 different teaching fellows' sections. The researcher knows that test scores are left-skewed within each section. Propose 2 different analysis approaches to address this professor's hypotheses. (6%)
II. A quadratic regression model to predict national test scores (Y) from parents' income (X) was calculated to be

$$\hat{Y} = 600 + 5X - 0.25X^2$$

Determine what value of income leads to the maximum estimated test score, and determine this maximum estimated test score. (4%)

III. A regression is run in order to determine whether the last n = 102 monthly returns of TSMC's Stock prices mimic that of UMC's Stock prices. Based on the output below, determine whether a slope of

(1) $\beta_1 = 0$ is reasonable. (2%)

(2) $\beta_1 = 1$ is reasonable. (4%)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.000278	0.006720	-0.041	0.967
UMC	0.783439	0.160266	4.888	3.97e-06

IV. Let $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ be two random samples from normal distributions with common variance σ^2. The pooled estimator of σ^2 is defined as

$$s_p^2 = \frac{(n_1-1)s_x^2+(n_2-1)s_Y^2}{n_1+n_2-2},$$

where $s_x^2 = \sum_{i=1}^{n_1}(X_i - \bar{X})^2/(n_1 - 1)$ and $s_Y^2 = \sum_{i=1}^{n_2}(Y_i - \bar{Y})^2/(n_2 - 1)$. Please show that s_p^2 is an unbiased estimator of σ^2. (5%)

3. A researcher perform a regression analysis to study the relationship between the prestige (Y) and the proportion of women (X) in 102 occupations. The results are obtained by performing analysis with the statistical software.

```
Call:
lm(formula = Y ~ X)
Residuals:
        Min       1Q      Median    3Q       Max
      -33.444   -12.391   -4.126   13.034   39.185
Coefficients:
              Estimate    Std. Error   t value   Pr(>|t|)
(Intercept)    48.693      2.30760     21.101    <0.001
X              -0.064      0.05        (a)____
---

Residual standard error: 17.17 on 100 degrees of freedom
Multiple R-squared:   0.014,  Adjusted R-squared:   0.004143
F-statistic: (b)______   on 1 and 100 DF,   p-value: 0.2362
```

(I) What is the correlation between prestige and the proportion of women? (3%)

(II) Please calculate the values in the blanks (a) and (b) above. (6%)

(III) Is there any evidence of a relationship between prestige and the proportion of women in occupations? (3%)

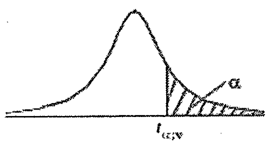
4. Based on the illustration for the two-sample t test below, answer the following questions (fill in the blank). (12%) （請在答案卷上作答）

	Two-sample t test	Paired t-test	McNemar's exact test	Friedman test
Q1) What outcome variable is the test appropriate for? (A) Continuous (B) Categorical (C) time to event	A	(1)_____	(2)_____	(3)_____
Q2) How many groups can the test handle? (A) One group (B) Two groups only (C) ≥3 groups	B	(4)_____	(5)_____	(6)_____
Q3) Is the data: (A)Independent (B) Correlated (C) Time-to-event	A	(7)_____	(8)_____	(9)_____
Q4) Do the data need to come from a normal distribution? (Yes / No)	Yes	(10)______	(11)_____	(12)_____

Numbers in the table represent $p(X=x)$ for a binomial distribution with n trials and probability of success p.

n	x	p = 0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75
8	0	0.430	0.168	0.100	0.058	0.017	0.004	0.001	0.000	0.000
	1	0.383	0.336	0.267	0.198	0.090	0.031	0.008	0.001	0.000
	2	0.149	0.294	0.311	0.296	0.209	0.109	0.041	0.010	0.004
	3	0.033	0.147	0.208	0.254	0.279	0.219	0.124	0.047	0.023
	4	0.005	0.046	0.087	0.136	0.232	0.273	0.232	0.136	0.087
	5	0.000	0.009	0.023	0.047	0.124	0.219	0.279	0.254	0.208
	6	0.000	0.001	0.004	0.010	0.041	0.109	0.209	0.296	0.311
	7	0.000	0.000	0.000	0.001	0.008	0.031	0.090	0.198	0.267
	8	0.000	0.000	0.000	0.000	0.001	0.004	0.017	0.058	0.100
9	0	0.387	0.134	0.075	0.040	0.010	0.002	0.000	0.000	0.000
	1	0.387	0.302	0.225	0.156	0.060	0.018	0.004	0.000	0.000
	2	0.172	0.302	0.300	0.267	0.161	0.070	0.021	0.004	0.001
	3	0.045	0.176	0.234	0.267	0.251	0.164	0.074	0.021	0.009
	4	0.007	0.066	0.117	0.172	0.251	0.246	0.167	0.074	0.039
	5	0.001	0.017	0.039	0.074	0.167	0.246	0.251	0.172	0.117
	6	0.000	0.003	0.009	0.021	0.074	0.164	0.251	0.267	0.234
	7	0.000	0.000	0.001	0.004	0.021	0.070	0.161	0.267	0.300
	8	0.000	0.000	0.000	0.000	0.004	0.018	0.060	0.156	0.225
	9	0.000	0.000	0.000	0.000	0.000	0.002	0.010	0.040	0.075
10	0	0.349	0.107	0.056	0.028	0.006	0.001	0.000	0.000	0.000
	1	0.387	0.268	0.188	0.121	0.040	0.010	0.002	0.000	0.000
	2	0.194	0.302	0.282	0.233	0.121	0.044	0.011	0.001	0.000
	3	0.057	0.201	0.250	0.267	0.215	0.117	0.042	0.009	0.003
	4	0.011	0.088	0.146	0.200	0.251	0.205	0.111	0.037	0.016
	5	0.001	0.026	0.058	0.103	0.201	0.246	0.201	0.103	0.058
	6	0.000	0.006	0.016	0.037	0.111	0.205	0.251	0.200	0.146
	7	0.000	0.001	0.003	0.009	0.042	0.117	0.215	0.267	0.250
	8	0.000	0.000	0.000	0.001	0.011	0.044	0.121	0.233	0.282
	9	0.000	0.000	0.000	0.000	0.002	0.010	0.040	0.121	0.188
	10	0.000	0.000	0.000	0.000	0.000	0.001	0.006	0.028	0.056

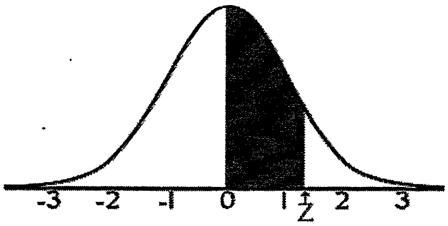
Table of the Student's *t*-distribution



The table gives the values of $t_{\alpha;v}$ where $\Pr(T_v > t_{\alpha;v}) = \alpha$, with v degrees of freedom

v \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587

Table for standard normal



Z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981