PART I. Multiple Choice Questions, 3 points each. You do not need to show detailed steps，無需列出計算過程。

1. How many tissues should a package of tissues contain? Researchers have determined that a person uses an average of 69 tissues during a cold. Suppose a random sample of 10,000 people yielded the following data on the number of tissues used during a cold: $\bar{x} = 54, s = 18$, We want to test the alternative hypothesis $H_a: \mu < 69$. Using the sample information provided, calculate the value of the test statistic for the test and state the rejection region for $\alpha = 0.05$.
   A) $z = -8.3333$, Reject H_0 if $z < -1.645$.　B) $z = -0.8333$, Reject H_0 if $z < -1.645$.
   C) $z = -8.3333$, Reject H_0 if $z < -1.96$.　D) $z = -0.8333$, Reject H_0 if $z < -1.96$.

2. A bottling company produces bottles that hold 12 ounces of liquid. Periodically, the company gets complaints that their bottles are not holding enough liquid. To test this claim, the bottling company randomly samples 81 bottles and finds the average amount of liquid held by the bottles is 11.9155 ounces with a standard deviation of 0.45 ounce. What is the *p*-value of this test and the conclusion at $\alpha = 0.05$?
   A) p = 0.4247, fail to reject the null hypothesis.　B) p = 0.0455, reject the null hypothesis.
   C) p = 0.4247, reject the null hypothesis.　D) p = 0.0455, fail to reject the null hypothesis.

3. Suppose a large labor union wishes to estimate the mean number of hours per month a union member is absent from work. The union decides to sample 20 of its members at random and monitor the working time of each of them for 1 month. At the end of the month, the total number of hours absent from work is recorded for each employee. If the mean and standard deviation of the sample are $\bar{x} = 9.9$ hours and s = 2.4 hours, find a 95% confidence interval for the true mean number of hours a union member is absent per month.
   A) 9.9 ± 0.2352　B) 9.9 ± 0.2512　C) 9.9 ± 1.05185　D) 9.9 ± 1.1232

4. A marketing research company is estimating the average total compensation of CEOs in the service industry. Data were randomly collected from 18 CEOs and the 95% confidence interval for the mean was calculated to be ($2,081,260, $5,243,002). Explain what the phrase "95% confident" means.
   A) 95% of the similarly constructed intervals would contain the value of the sample mean.
   B) 95 %of the sample means from similar samples fall within the interval.
   C) In repeated sampling, 95% of the intervals constructed would contain μ.
   D) 95% of the population values will fall within the interval.

5. The university police department must write, on average, five tickets per day to keep department revenues at budgeted levels. Suppose the number of tickets written per day follows a Poisson distribution with a mean of 8.9. What is the probability that at least five tickets are written on a randomly selected day?
   A) 0.0635　B) 0.0584　C) 0.9772　D) 0.9416

6. Suppose the candidate pool for two appointed positions includes 6 women and 9 men. All candidates were told that the positions were randomly filled. What is the probability that two women are selected to fill the appointed positions?
   A) 0.3429　B) 0.1601　C) 0.3605　D) 0.1429

7. A random variable follows a binomial probability distribution with the probability of success p = 0.2. For a sample of 100 trials, what is the approximated probability of 15 or fewer successes using normal probability distribution?
   A) 0.2643　B) 0.0853　C) 0.1292　D) 0.0548

8. Given below is a bivariate distribution for the random variables x and y.

f(x,y)	x	y
0.2	50	80
0.5	30	50
0.3	40	60

Compute the variance of $x+y$, *i.e.*, $Var(x+y)$, which is

A) 259 B) 364 C) 312 D) 415

9. A bank reviewed its credit card policy with the intention of recalling some of its credit cards. In the past approximately 5% of cardholders defaulted, leaving the bank unable to collect the outstanding balance. The bank also found that the probability of missing a monthly payment is 0.2 for customers who do not default. Of course, the probability of missing a monthly payment for those who default is 1. Given that a customer missed one monthly payment, what is the probability that the customer will default?

A) 0.2083 B) 0.3412 C) 0.1554 D) 0.4355

10. A department of transportation's study on driving speed and miles per gallon for heavy-duty trucks resulted in the following data.

Speed (Miles per hour)	50	50	40	60	30
Miles per gallon	4	6	11	3	16

Compute the sample correlation coefficient, which is equal to.

A) -0.8512 B) -0.7445 C) -0.2358 D) -0.9689

Use the following information to answer questions 11 and 12

A researcher surveyed 100 people regarding whether using ATM to withdraw their cash is associated with their ages. The survey result is shown as below：

	Use ATM	Do not use ATM
Younger	20	30
Elder	15	35

11. Use the two proportion equality z test to test whether the proportions of using ATM are equal between younger population and elder population? What is the p value of this test？

A) $p>10\%$ B) $5\%<p<10\%$ C) $1\%<p<5\%$ D) $p<1\%$

12. Use the independence chi-square test to test whether age and the use of ATM are mutually independent. The calculated χ^2 statistic is

A) Greater than 5 B) Between 5 and 3 C) Between 3 and 1 D) Less than 1

Use the following information to answer questions 13 and 14

A bank's manager likes to know whether the average mutual funds sold per month are equal between branches (office A, B and C) over past years. He conducts an ANOVA study and obtains the following results：

$\overline{X}_A = 31.25, \overline{X}_B = 41, \overline{X}_C = 46$ are group means and $n_A = 4, n_B = 5, n_C = 6$ are the number of repeated measure, $S_A^2 = 48.25, S_B^2 = 28.5, S_C^2 = 68.8$, where $S_j^2 = \sum_j (X_{ij} - \overline{X}_j)^2 / (n_j - 1)$, are the sample variance for office A, B and C.

13. The manager first likes to test whether the average monthly fund sold are equal between its three branch offices (i.e., whether $\mu_A = \mu_B = \mu_C$). The calculated F statistic is
A) Less than 2　B) Between 2 and 4　C) Between 4 and 6　D) greater than 6

14. The manger likes to specifically test whether the average number between office A and office B are equal (i.e., $H_0: \mu_A = \mu_B$) by using the *Bonferroni* multiple comparison method. The p value of this test is approximately equal to
A) p>10%　B) 5%<p<10%　C) 1%<p<5%　D) p<1%

15. Consider a simple regression model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. There are four observations (x, y), in which the value of x_i (i=1,..,4) are known as -3, -1, 1, 3, respectively. It is known that the least square estimator of slope coefficient can be expressed in terms of the weighted average of observed y_i, that is, $\hat{\beta}_{LS} = \sum_i w_i y_i$.
Find the value of (w_1, w_2), which is
A) (1/8, 3/8)　B) (3/8, 1/8)　C) (3/20, 1/20)　D) (3/20, 1/20)

16. Consider a simple log-linear model $\ln y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, what is the exact expected percentage change of y when there is one unit of x change?
A) 100(expβ_1)　B) 100(expβ_1-1)　C) 100(lnβ_1)　D) 100(lnβ_1-1)

17. We try to fit a regression model without intercept, i.e., $y_i = \beta x_i + \varepsilon_i$ by using the following three (x, y) observations, (1.0, 1.0), (4.5, 9.0), (7.0, 20.1). We also know the heterocedasticity problem is existent in a manner of $Var(\varepsilon_i) = \sigma^2 x_i$. The weighted least-square estimate (*WLS*) of β is
A) Less than 3　B) Between 3 and 4　C) Between 4 and 5　D) greater than 5

18. Which of the following statements regarding the least square estimator of a simple regression $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ is mostly <u>incorrect</u>?
A) If the mean of X is positive, then an overestimate of β_0 can lead to an underestimate of β
B) The precision of the slope estimator decreases as the variation of the X increase
C) The lower ratio of the standard error of the regression (s) to the mean of Y, the more closely the data fit the regression line.
D) The residual variance s^2 is an unbiased and consistent estimator of the error variance σ_ε^2.

19. For a multiple regression model $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$. You are given that

	Partial correlation	Standardized coefficient	Elasticity
X_1	0.64	0.50	0.20
X_2	-0.04	-0.01	-0.01
X_3	0.70	0.40	0.60

Which of the following is implied by the above results?
A) 16% of the variance of Y not accounted by X_1 and X_2 is accounted for X_3
B) An increase of 1 standard deviation in X_1 will lead to an increase of 0.64 standard deviation in Y
C) An increase of 1% in X_1 will lead to an increase of 0.20% in Y
D) An increase of one unit in X_2 will lead to a decrease of 0.04 units in Y

20. Which of the following statements is <u>incorrect</u>?
A) The least square estimator is essentially a sample estimator using the method of moment
B) A necessary condition for a consistent regression estimator is the independence between error term and

explanatory variables

C) Regress the least square residual on the regressors can lead to an zero R-square.

D) The least square estimator is efficient if the error terms are not mutually independent.

PART II. You <u>must show detailed steps</u> for each question to get credits. (需列出計算過程才能獲得分數).

1. Consider the following hypothesis test.

   $H_o: \mu = 40$ vs. $H_a: \mu \neq 40$

   a. (5 points) Use $\alpha = 0.05$. A sample of 64 provides a sample mean of 41 and a sample standard deviation of 4. What is the power of the statistical test when the actual population mean is 42?

   b. (5 points) Use $\alpha = 0.05$. If the population standard deviation is 10, how large a sample should be taken if the researcher is willing to accept a 0.05 probability of making a Type II error when the actual population mean is 42?

2. A congressional committee has been charged with conducting a sample survey to obtain information about the percentage of people not covered by health care insurance.

   a. (5 points) A simple random sample of 4000 individuals provides 1000 responses of no health care insurance. Compute the 95% confidence interval for the population proportion of people not covered by health care insurance.

   b. (5 points) What sample size would you recommend if the committee's goal is to estimate the proportion of individuals without heal care insurance with a margin of error of 0.03? Use a 95% confidence level.

3. (12 points) Consider an one-way ANOVA with k treatments, show that the expectation of mean square treatment (MST) is equal to the common variance, i.e., $E(MST) = E\left(\frac{\sum_{j=1}^{k} n_j(\bar{X}_j - \bar{\bar{X}})}{k-1}\right) = \sigma^2$ only if treatment variances are equal, ($\sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 = \sigma^2$), and all treatment population means are equal, ($\mu_1 = \mu_2 = ... = \mu_k = \mu$), where $\bar{X}_j, n_j$ are the mean and the repeated number for treatment j. $\bar{\bar{X}}$ is the grand mean. (*Hint: you may use the fact that* $E(\bar{\bar{X}})^2 = (\sigma^2 / n) + \mu^2$), $E(\bar{X}_j)^2 = (\sum_j \sigma_j^2 / n_j) + \mu_j^2$).

4. Consider a single-factor ANOVA, in which the factor has 3 levels, the response variable y_{ij} can be expressed by the following additive terms:

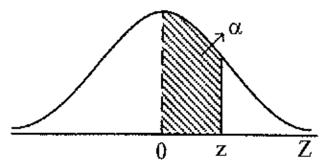
   $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, where μ is the grand population mean, α_i is the effect of factor level i on y_{ij}, i=1,2,3, and ε_{ij} is the error term. This model suggests that the ANOVA problem can be tested by a (linear) regression model with dummy variables.

   (a). (4 points) Show the regression model and define the dummy variables that can reflect all factor effects.

   (b). (4 points) Show the testable regression hypothesis that is equivalent to test whether all factor means are equal (i.e., $\mu_1 = \mu_2 = \mu_3$, where μ_i is the factor mean for level i) in the ANOVA problem.

附表1　z分配表

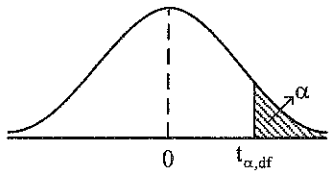
$P(0<Z<z)=\alpha$



z	0.00	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
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998

附表2　t分配臨界值表

$P(t_{df}>t_{\alpha,df})=\alpha$



df	α 0.25	0.20	0.15	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.001	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.710	15.890	31.820	63.660	127.30	318.30	636.60
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.090	22.330	31.600
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.210	12.920
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291