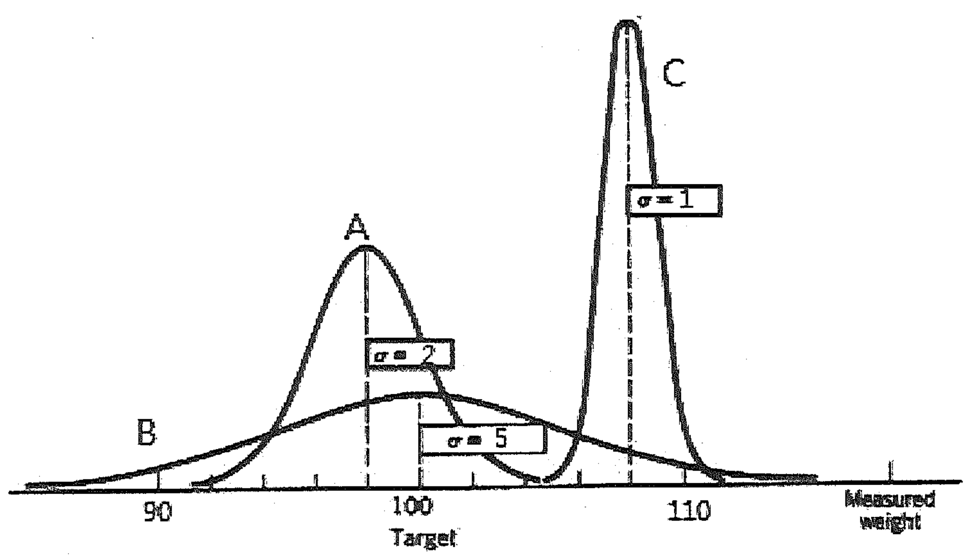
國立臺北大學 104 學年度碩士班一般入學考試試題

系（所）組別：統計學系
科　　　目：統計學

☑可　☐不可使用計算機

I. (50%)

1.1 (10%) An estimator that has small variance (but may be biased) is called precise. An estimator that has small MSE (Mean squared error) is called accurate. To illustrate: A standard 100-gm mass was weighted many, many times, on scale A, and the distribution of measurements is graphed below. Similarly, on scale B, and finally on scale C.



a. Calculate MSE for scales A, B, and C. (5%)

b. Which scale is most precise? Which scale is most accurate? Which is more important for an estimator to be precise or accurate? Why? (5%)

1.2. (20%) On the average, 6 cars arrive at the drive-up window of a bank every hour. Define the random variable X to be the number of cars arriving in any hour.

a. What is the appropriate probability distribution for X? Explain how X satisfies the properties of the distribution. (5%)

b. Find the mean and variance of *X*. (5%)

c. Define the random variable Y to be the time between arrivals of customers at the drive-up window of a bank. Write the probability density function for Y. (5%)

d. Find the mean and variance of Y. (5%)

1.3. (10%) An investor has decided to form a portfolio by investing 25% of his money into AA stock and 75% into BB stock. The following data are the expected returns and standard deviations of the investments.

Investment	Expected return	Standard deviation
AA stock	8%	12%
BB stock	15%	22%

a. Find the expected return on the portfolio. (5%)

b. If ρ = 1, .5, 0 find the standard deviation of the portfolio. Explain the relationship between the standard deviation of the portfolio and the correlation coefficient (ρ). (5%)

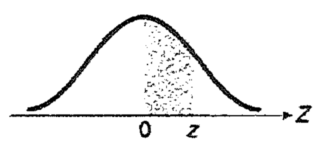
試題隨卷繳交　　　　接背面

1.4. (10%) The chairperson of department of statistics in a certain college believes that 70% of the department's graduate assistantships are given to international students. A random sample of 100 graduate assistants is taken.

a. What is the probability that the sample proportion $\bar{p}$ will be within $\pm .05$ of the population proportion p? (5%)

b. If the chairperson would like to reduce the sampling error, how large a sample should be drawn to estimate with 95% confidence the population proportion to be within 0.05? (Using the assumption p =.70., margin of error = 0.05) (5%)

Areas of the standard normal distribution. The entries in this table are the probabilities that a standard normal random variable is between 0 and z (the shaded area).



Second Decimal Place in z

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									
4.0	0.49997									
4.5	0.499997									
5.0	0.4999997									

II. For the following questions, use 0.05 as level of significance unless otherwise specified.

2.1. (15%) A survey finds that 200 of lower-income households have Internet access at home as compared to 280 of upper-income households. Assume that the data are based on random samples of size 400 each.

a. Does this demonstrate that upper-income households are <u>more likely</u> to have Internet access than the lower -income households? Conduct an appropriate test to answer the question. (10%)

b. Construct a 95% confidence interval for the difference of the proportions. (5%)

2.2. (10%) The following data are from a completely randomized design:

	Treatment		
Subject	A	B	C
1	32	44	33
2	30	43	36
3	30	44	35
4	26	46	36
5	32	48	40

Can you reject the null hypothesis that the means of the three treatments are equal?

2.3 (15%, 5% each) The following data are the monthly salaries (Y, in dollars) and the grade point averages (GPA, X) for students who obtained a bachelor's degree in business administration.

students	GPA	Monthly Salary ($)
1	2	3600
2	3.2	3900
3	3.4	4400
4	3	3800
5	3.3	4400
6	3.1	3900

The following results are obtained: MSE(mean sqaure error) $= 53750$, $\sum\left(x_i - \bar{x}\right)^2 = 1.3$

The estimated regression equation for these data is $\hat{Y} = 2500 + 500X$

a. Develop a point estimate of the starting salary for a student with a GPA of 3.0.

b. Develop a 95% confidence interval for the mean starting salary for all students with a 3.0 GPA.

c. Develop a 95% prediction interval for Ryan Dailey, a student with a GPA of 3.0.

2.4. (10%) Suppose the serum cholesterol levels for all 20- to 25-year-old males of the population is normally distributed with a mean of 180 mg/100ml and the standard deviation is 40 mg/100ml. We would expect that the mean cholesterol level of a special diet group to be higher than 180 mg/100ml. (Assume that the cholesterol levels are normally distributed and have the same standard deviation, i.e., 40 mg/100ml.) To test whether the mean cholesterol level of this special diet group to be higher than 180 mg/100ml, we conduct a hypotheses. If we use a sample of size 100 and assume the true mean of this special diet group is 196 mg/100ml, what is the power of the test using level of significance 0.025?

國立臺北大學 104 學年度碩士班一般入學考試試題

系（所）組別：統計學系

科　　　目：統計學　　　　　　　　　　　　　　　　第 4 頁　共 4 頁

☑可　☐不可使用計算機

Table 1. Entries in the table give $F_{0.05}$ critical values, where 0.05 is the area or probability in the upper tail of the F distribution. For example, with 4 numerator degrees of freedom, 8 denominator degrees of freedom, and a .05 area in the upper tail, $F_{0.05}$= 3.84.

F critical values													
		numerator											
		1	2	3	4	5	6	7	8	9	10	11	12
denominator	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48

Table 2. Percentiles of the t distributions: Entry is $t(A;V)$ where $P\{t(v) \le t(A;v)\} = A$ for example, $P\{t(1) \le t(12.706,1)\} = 0.975$,v is the degree of freedom.

	A				A	
v	**0.975**	**0.95**		v	**0.975**	**0.95**
1	12.706	6.314		7	2.365	1.895
2	4.303	2.92		8	2.306	1.86
3	3.182	2.353		9	2.262	1.833
4	2.776	2.132		10	2.228	1.812
5	2.571	2.015		11	2.201	1.796
6	2.447	1.943		12	2.179	1.782

試題隨卷繳交