所別：統計研究所碩士班 一般生 科目：數理統計

EXAM ON MATHEMATICAL STATISTICS

No calaulator is allowed. The first six problems are multiple choice problem. Find out all the statements that are correct. Please treat each of the statements as a yes-no problem. If your answer is correct , then you get all the points. If you make four correct judgements out of five, you get half of the points. Otherwise you get zero point. There is no penalty for incorrect answers. For example, if the correct answer is "abd" but your answer is "abcd" for a problem with 10 points, you get 5 points. You don't need to do any justification for problems one to six. In order to get a full credit for the seventh problem, you need to write down your calculation in order to justify your answers. The probability distribution function of the standard normal distribution is on the last page.

1. (10%)Let X_1, X_2,...... X_i,......be a sequence of i.i.d. random variable with a common distribution whose probability density function is symmetric about y−axis. $\frac{1}{\sqrt{n}}(X_1 + X_2 + \ldots\ldots X_n)$ converges to a normal distribution as $n \to \infty$. Which of the following can be the probability density function of X_1?

   (a) $\frac{1}{\pi}\frac{1}{1+x^2}$,for all $x \in R$

   (b) $\frac{1}{2x^2}$, for $|x| > 1$.

   (c) $\frac{1}{4}$ for $|x| \le 2$.

   (d) $e^{-2|x|}$ for all $x \in R$.

   (e) $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$for all $x \in R$.

2. (10%) Let X_1, X_2, X_3 be i.i.d. Poisson random variable with parameter λ, while Y_1, Y_2, Y_3 are i.i.d. gamma random variables with parameters (α, β). Let X_4 be a Poisson random variable with parameter 2λ. Which of the following statements MUST be true?

   (a) The statement "$Cov(X_1, X_4) = 2\lambda$" may be true.

   (b) $E(X_1 + X_4) = 3\lambda$.

   (c) $P(X_1 < X_2 < X_3) = \frac{1}{6}$.

   (d) $P(Y_1 < Y_2 < Y_3) = \frac{1}{3}$.

   (e) $X_2 + X_4$ is a Poisson random variable.

3. (10%) Let X_1, X_2,...... X_{10} be a sequence of i.i.d. random variables that follow a normal distribution $N(\mu, \sigma^2)$. We would like to make inference about the mean μ. Let s be the sample standard deviation

$$s = \sqrt{\frac{1}{10-1}\sum_{i=1}^{10}(X_i - \overline{X})^2}.$$

   (a) If the variance σ^2 of the population is known, then the shortest 95% confidence interval for μ we can obtain is $\overline{X} \pm z\frac{\sigma}{\sqrt{10}}$, where z is the 95th percentile of the standard normal distribution.

注意：背面有試題

所別：統計研究所碩士班 一般生 科目：數理統計

(b) If the variance σ^2 of the population is unknown, then the shortest confidence interval we can obtain in that case is symmetric around the sample mean $\overline{X}$.

(c) The shortest confidence interval obtained without the knowledge of variance is longer than the shortest confidence interval obtained with the knowledge of variance. (We consider the same level of confidence for both cases.)

(d) If the sample size is very large, then the shortest confidence interval obtained without the knowledge of variance is approximately the same as the shortest confidence interval obtained with the knowledge of variance. (We consider the same level of confidence for both cases.)

(e) If the variance σ^2 of the population is unknown, we can obtain a shorter confidence interval in the presence of more observations.

4. (16%) A communication device transmits sequences of binary digits, 0 or 1. A transmission error occurs if a digit is sent as zero but received as a one, or vice versa. Assume that the probability of error of each digit transmitted is p,and all transmission errors are independent (that is, the probability of an error occurring at a given digit is independent of the outcomes of all other digits). To ensure accuracy of a message, the same sequence of n binary digits is transmitted twice. The receiver notes the number of digits that differ in the two received messages. For instance, with $n = 8$ the two received message might be 00110000 and 00100001 , so $X = 2$. The fourth and the eighth digits differ in the two message. Which of the following statements must be true?

(a) The probability that the third digit differs in the two message is $p(1-p)$. (In this case, $n \geq 3$.)

(b) The probability that exactly x digits differ in the two message is $\binom{n}{x}[2p(1-p)]^x[1-2p(1-p)]^{n-x}$.

(c) When we observe a message with n digits, and X of them are different. We would like to estimate p by maximum likelihood estimation. The problem can be reduced to a problem that maximizes $\log\binom{n}{X}+X\log[2p(1-p)]+(n-X)\log[1-2p(1-p)]$.

(d) When we observe a message with n digits, and X of them are different. We would like to estimate p by maximum likelihood estimation. The maximum likelihood estimator for p is $\frac{X}{n}$.

(e) If a large proportion of digits are different in the two messages, then it is very likely that p is close to 1.

5. (16%) Six tests are done for the two hypotheses $H_0 : \mu \leq 0$, versus $H_1 : \mu > 0$ and $H_0 : \mu = 0$, versus $H_1 : \mu \neq 0$ when samples are drawn from a normal distribution with a known variance $\sigma^2 = 100$. The plots of power functions are as in Figure 1. The six tests are

A $H_0 : \mu \leq 0$, versus $H_1 : \mu > 0$, sample size $n = 1$.

B $H_0 : \mu \leq 0$, versus $H_1 : \mu > 0$, sample size $n = 4$.

注意：背面有試題

C $H_0 : \mu \leq 0$, versus $H_1 : \mu > 0$, sample size $n = 100$.

D $H_0 : \mu = 0$, versus $H_1 : \mu \neq 0$, sample size $n = 1$.

E $H_0 : \mu = 0$, versus $H_1 : \mu \neq 0$, sample size $n = 4$.

F $H_0 : \mu = 0$, versus $H_1 : \mu \neq 0$, sample size $n = 100$.

Which of the following are correct?

(a) Plot 1 corresponds to test A.
(b) Plot 5 corresponds to test B.
(c) Plot 2 corresponds to test E.
(d) Plot 3 corresponds to test D.
(e) Plot 6 corresponds to test F.

6. (16%) Consider two hypothesis testing problems:

A : Let X be a normal random variable with n trials and the probability p of success. $H_0 : p = 0.5$ versus $H_1 : p \neq 0.5$.

B : Let $X_1, X_2, \ldots\ldots X_n$ be a random sample from an exponential distribution with the density function $f(x|\theta) = \theta \exp[-\theta x]$. $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

We would like to test each of the hypothesis by a generalized likelihood ratio test procedure.

(a) Test A: The generalized likelihood ratio always rejects for large values of $|X - \frac{n}{2}|$.

(b) Test A: If $n = 10$, then the significant level of the tests that rejects $|X - \frac{n}{2}| > 2$ is between 0.08 and 0.12.

(c) Test A: If $n = 100$, then the significant level of the tests that rejects $|X - \frac{n}{2}| > 10$ is less than 0.06. (Hint: Approximate by normal distribution)

(d) Test B: The rejection region is in the form $\{\overline{X} \exp[-\theta_0 \overline{X}] \leq c\}$ for some constant c.

(e) The plots of the power function for both tests are symmetric.

7. Let $X_1, X_2, \ldots\ldots X_n$ be a be a sequence of i.i.d. random variable with a common probability density function

$$f(x|\theta) = \theta x^{\theta-1}, 0 \leq x \leq 1, 0 < \theta < \infty.$$

(a) (16%)Find $\widehat{\theta}$, the MLE (maximum likelihood estimator) of θ and the variance of $(\widehat{\theta})^{-1}$.

(b) (6%)Find the method of moment estimator for θ .

注意：背面有試題

所別：統計研究所碩士班 一般生 科目：數理統計

Figure 1: Plots of power functions for six hypothesis testing problems

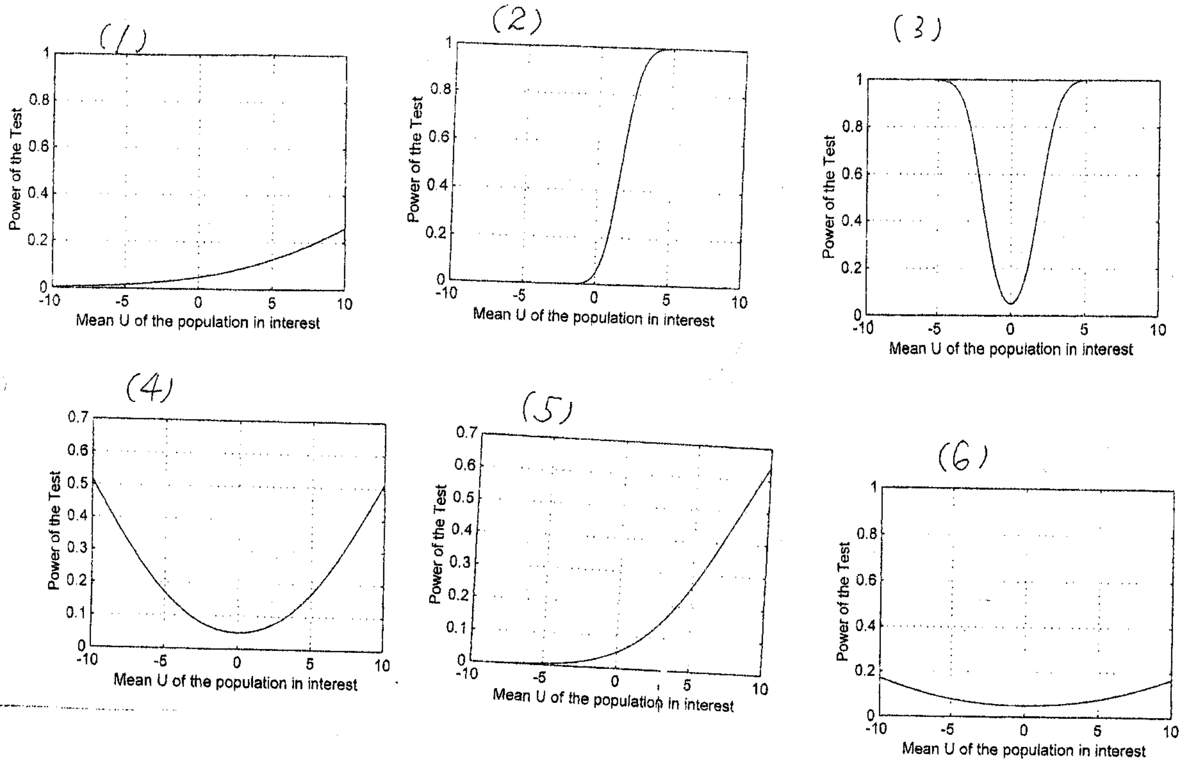

Table entry for z is the area under the standard normal curve to the left of z.

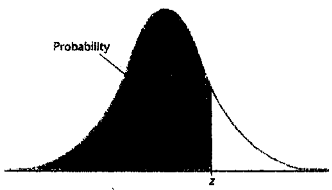

TABLE A Standard normal probabilities (*continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998