MATH372301

This question paper consists of 7 printed pages, each of which is identified by the reference **MATH3723**.

Only approved basic scientific calculators may be used. Statistical tables are provided on p.7.

Examination for the Module MATH3723
(May/June 2005)

Statistical Theory

Time allowed: **3 hours**

Do not answer more than four questions.
All questions carry equal marks.

1. (a) Define the density of the Gaussian distribution $\mathcal{N}(a, \sigma^2)$.

Let $X = Z^2$ where $Z \sim \mathcal{N}(0, \sigma^2)$. Using the rule for computing the density, $f_X(x, \sigma^2) = dP(X \le x)/dx$, or otherwise, show that the density $f_X(x, \sigma^2)$ can be represented in the form,

$$f_X(x, \sigma^2) = x^{-1/2}(2\pi\sigma^2)^{-1/2} e^{-x/2\sigma^2}, \quad x \ge 0.$$

For the sample $X = (X_1, \ldots, X_n)$ write down the likelihood function.

Using this likelihood function or otherwise, show that the statistic $T(X_1, \ldots, X_n) = \sum_{i=1}^n X_i$ based on observations $X_1, \ldots, X_n$, is sufficient for σ^2.

(b) Define what is meant by a best estimator and a best linear estimator in the mean square error sense, for a parametric density family $\{f(x, \theta), \theta \in \Theta\}$.

If both estimators exist for some general parametric family, which one has the smallest Mean Square Error and which one the largest?

For the density family $\{f_X(x, \sigma^2)\}$ from part (a), given the sample $X = (X_1, \ldots, X_n)$, find the Maximum Likelihood Estimator of σ^2. Is it unbiased?

 QUESTION 1 CONTINUED ...

(c) Define the notion of Fisher's information for a parametric density family.

For the family of densities $\{f_X(x, \sigma^2)\}$ from part (a), compute the Cramér-Rao lower bound for the unbiased estimators of the parameter.

Is this bound attained by any unbiased estimator? Justify your answer.

Hint: if necessary you may use the values of the even moments of the standard Gaussian distribution $\mathcal{N}(0, \sigma^2)$, the moment of the order $2k$ equals $\sigma^{2k}(2k-1)!!$, where $n!!$ is the product of all natural numbers from one to n of the same parity with n, that is,

$$n!! = \begin{cases} n \times (n-2) \times \ldots 1, & n \text{ odd}, \\ n \times (n-2) \times \ldots 2, & n \text{ even}. \end{cases}$$

2. (a) Define the notion of the Method of Moments Estimator and the Maximum Likelihood Estimator for a 1-parametric statistical distribution.

Define the uniform distribution $U[0, 2\theta]$ for $\theta > 0$ via its density.

For the sample $X = (X_1, \ldots, X_n)$ from this distributions find the estimator of θ based on Method of Moments. Is it unbiased? Compute its Mean Square Error.

(b) For the distribution family from part (a) and for the sample $X = (X_1, \ldots, X_n)$ find the density of the order statistic $X_{(n)} = \max(X_1, \ldots, X_n)$.

Find a Maximum Likelihood Estimator of θ. Is it unbiased? Compute its Mean Square Error.

(c) For one observation $X = X_1$ from the distribution in part (a), given a Bayesian prior $\theta = 1/2$ and $\theta = 1$ with probabilities $1/2$, compute the Bayesian Estimator for θ. Compute the Bayesian Mean Square Error of this estimator.

CONTINUED...

3. (a) Define the notion of sufficient statistics.

State the Factorization Criterion for sufficient statistics and the Rao-Blackwell Theorem on improvement of unbiased estimators.

Describe the notion of confidence interval or set for an unknown parameter.

Given a sample $X_1, \ldots X_n$ from a distribution family $U[\theta - 1/2, \theta + 1/2]$ with unknown location θ, suggest a confidence interval (CI) for $\theta \in R$ with confidence 0.99, either exact or approximate.

Hint: for an approximate CI, normal approximations could be used; for an exact one, one could, for example, find an appropriate sufficient statistic (T_1, T_2) and a distribution of either of the T_i's.

(b) A sample $X_1, \ldots, X_n$ is given from the distribution family $U[\theta - 1/2, \theta + 1/2]$ with unknown location $\theta \in R$, and Bayesian prior $\theta \sim U[0, 1]$. Give the definition of credible interval with given credibility $1 - \alpha$ for the unknown parameter $\theta \in R$, and construct a credible set for $\theta \in R$ with credibility $1 - \alpha$ for $n = 1$.

(c) In the model from (a) and for the sample $X = (X_1, \ldots, X_n)$ with sample size $n \geq 2$, find the density of the order statistic $X_{(1)} = \min(X_1, \ldots, X_n)$ and $X_{(n)} = \max(X_1, \ldots, X_n)$.

Find a sufficient statistic of dimension two.

Suggest an unbiased estimator of $\theta \in R$ as a linear function of this statistic. Show that this estimator is, indeed, unbiased.

4. (a) Define the notion of Bayesian Estimator with the Mean Square Error criterion.

Consider a sample $X = (X_1, \dots, X_n)$ from the Bernoulli distribution, $\theta^x(1 - \theta)^{1-x}$, $x = 0, 1,\ 0 \le \theta \le 1$. Let the prior for θ be $P(\theta = 0.9) = P(\theta = 0.1) = 1/2$. Compute the Bayesian Estimator of θ.

Hint: it may be useful to compute the two conditional posterior probabilities up to a normalization constant, and only then compute the latter constant.

(b) State the definition of the Bayesian decision rule in the hypothesis testing problem. For the distribution family from part (a) and with the same Bayesian prior, consider the hypotheses $H_0 : \ \theta = 0.1$ and $H_1 : \ \theta = 0.9$. Argue why the decision rule "accept H_1 iff $\bar{X} \ge 1/2$" is Bayesian.

Find a sample size n sufficient for the Bayesian error e_B to be at most 0.001.

State the theorem you use here about normal approximations; explain why all assumptions of this theorem are satisfied.

(c) For the Bernoulli distribution family from part (a), $\theta^x(1 - \theta)^{1-x}$, $x = 0, 1$, $0 \le \theta \le 1$, find the classical approximate critical region for testing hypotheses $H_0 : \theta = 1/2$ vs $H_1 : \theta > 1/2$ with significance level 0.05, based on observations $X = (X_1, \dots, X_n)$ and using the Neyman-Pearson Lemma.

5. (a) Define Bernoulli trials with probability of success p.

Write down the linear unbiased estimator of $p \in [0, 1]$, given one Bernoulli observation X_1.

Considering the class of all linear estimators αX_1, show that there exist biased estimators that are strictly better in the mean square sense than the unbiased one just found, assuming $1/3 \le p \le 2/3$.

(b) State the Rao-Blackwell Theorem on improvement of unbiased estimators.

Find a single sufficient statistic for the distribution from part (a) given a sample $X = (X_1, X_2, X_3)$.

Apply the Rao-Blackwell Theorem to the estimator X_1 and this sufficient statistic. [Hint: for example, show that $E(X_i \mid \bar{X}_3) = \bar{X}_3$, $\forall i$, without computation of conditional probabilities.]

Compute the Mean Square Error of the improved estimator.

(c) The sample mean value $\bar{X}_n$ from a sample $X = (X_1, \ldots, X_n)$ from the Bernoulli distribution $B(p)$ is given. Is this enough to construct the most powerful test for testing simple hypotheses about p? Justify your answer.

Using the Maximum Likelihood Ratio Test, and with the help of Wilks' theorem or otherwise, assuming n large, construct an approximate critical region of approximate size α for testing the hypothesis $H_0 : p = p_0$ vs. $H_1 : p \neq p_0$, where $0 < p_0 < 1$. What kind of table is needed?

END OF QUESTIONS

Normal distribution (areas)

Area ($\alpha = P(Z > z)$) in the tail of the standardized Normal curve, $Z \sim N(0,1)$, for different values of z. Example: Area beyond $z = 1.96$ (or below $z = -1.96$) is $\alpha = 0.02500$. For Normal curve with $\mu = 10$ and $\sigma = 2$, area beyond $x = 12$, say, is the same as the area beyond $z = \dfrac{x-\mu}{\sigma} = \dfrac{12-10}{2} = 1$, i.e. $\alpha = 0.15866$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
4.0	.00003	.00003	.00003	.00003	.00003	.00003	.00002	.00002	.00002	.00002

α	0.4	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.001
z_α	0.2533	0.6745	0.8416	1.0361	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902

END