Statistics Ph.D. Qualifying Exam: Part II

November 20, 2010

Student Name: ______________________________

1. Answer 8 out of 12 problems. Mark the problems you selected in the following table.

Problem	1	2	3	4	5	6	7	8	9	10	11	12
Selected												
Scores												

2. Write your answer right after each problem selected, attach more pages if necessary. **Do not** write your answers on the back.

3. Assemble your work in right order and in the original problem order. (Including the ones that you do not select)

4. You can use the N(0,1) distribution table as attached.

1. Let X, Y and Z be independent uniform(0,1) random variables.

   (a) Find the cdf of $W = XYZ$, $P(W \leq w)$.

   (b) Find the pdf of $W = XYZ$.

2. Let $X_1, ..., X_n$ be a random sample from an exponential distribution with mean 1. Let $Y_i = X_{(i)}$ be the i-th order statistics. Give the pdf of each of the following:

   (a) The smallest order statistic, Y_1.

   (b) The largest order statistic, Y_n.

   (c) The smallest and largest order statistics, Y_1 and Y_n.

   (d) The sample range, $R = Y_n - Y_1$.

3. (a) Let T_n, $n \geq 1$ be a sequence of random variables such that $n^{1/2}(T_n - \theta)$ converges in distribution to $N\big(0, \sigma^2(\theta)\big)$. If g is differentiable, find the asymptotic distribution of $g(T_n)$.

(b) Suppose that $X_1, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$. Find g such that the limiting distribution of $g(S_n^2)$ does not depend on σ^2, where $S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2/(n-1)$ and $\bar{X}_n = \sum_{i=1}^n X_i/n$.

4. Consider a linear model

$$\mathbf{y} = \mathbf{X}\beta + \epsilon,$$

where $\mathbf{y}$ and ϵ are $n \times 1$ random vectors, $\mathbf{X}$ is $n \times p$ matrix of full rank p, and β is a $p \times 1$ a vector of p unknown parameters. Assume that the mean and variance-covariance matrix of ϵ are $E(\epsilon) = 0$ and $V(\epsilon) = \sigma^2 I$.

(a) Derive the the least square estimator $\hat{\beta}$ of β.

(b) Derive variance-covariance matrix of $\hat{\beta}$.

(c) Explain why $\hat{\beta}$ is the best linear unbiased estimator of β.

5. Let $\{(X_{i,1}, \dots, X_{i,n}), i = 1, 2\}$ be independent random samples from normal distributions with means μ_i and variance σ^2 respectively. Let $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{i,j}, i = 1, 2$ and $S_i^2 = \sum_{j=1}^{n}(X_{i,j} - \bar{X}_i)^2, i = 1, 2$. Put $Y_i = \frac{\bar{X}_i}{\sqrt{\hat{\sigma}^2}}, i = 1, 2$, where $\hat{\sigma}^2 = (S_1^2 + S_2^2)/(2n - 2)$.

   (a) Obtain the joint pdf (probability density function) of $\{Y_1, Y_2\}$ under the assumption ($\mu_i = 0, i = 1, 2$).

   (b) What is the pdf of $Y_1 - Y_2$ under the assumption $\mu_1 = \mu_2$?

6. Let $X_1, \ldots, X_n$ be a random sample from a Normal $(\mu, 1/\tau)$ population. Assume the following prior specifications on μ and τ: $\mu|\tau \sim N(\mu_0, \frac{1}{\lambda\tau})$, $\tau \sim Gamma(\alpha, \beta)$.

   (a) Show that this prior specification is conjugate for this problem.

   (b) Find the posterior distribution of $\sqrt{\frac{\lambda\alpha}{\beta}}(\mu - \mu_0)$.

7. Let $Y_1, \ldots, Y_N$ be independent random variable such that $Y_i \sim \text{Binomial}(n_i, p_i)$, where

$$P_i = \frac{1}{1 + e^{\alpha + \beta x_i}},$$

and x_i is a fixed covariate, $i = 1, \ldots, N$.

(a) Find a set of jointly sufficient statistics for (α, β).

(b) Suppose $\tilde{\alpha}, \tilde{\beta}$ are the estimates of α and β that minimize

$$Q = \sum_{i=1}^{N} n_i \hat{p}_i (1 - \hat{p}_i)(\alpha + \beta x_i - l(Y_i/n_i))^2,$$

where $\hat{p}_i = Y_i/n_i$ and l_i is some function of $Y_i/n_i, i = 1, \ldots, N$. Find $\tilde{\alpha}$ and $\tilde{\beta}$.

(c) Let $\hat{\alpha}, \hat{\beta}$ be MLE's of α and β. Give an argument to show that

$$\text{Mean Square Error}(\hat{\alpha} + \hat{\beta}) \leq \text{Mean Square Error}(\tilde{\alpha} + \tilde{\beta})$$

8. Let $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ be independent samples from Exponential(λ) and Exponential(μ) populations respectively.

   (a) Construct a likelihood ratio test of

$$H_0 : \lambda = \mu \quad \text{versus} \quad H_1 : \lambda \neq \mu.$$

   (b) Give the critical values of this test in terms of percentiles of one of the standard distributions.

9. Let $X_1, \ldots, X_n$ be a random sample from the density $f(x, \theta) = \frac{\log \theta}{\theta - 1} \theta^x, 0 < x < 1, \theta > 1$.

(a) Obtain a sufficient and complete statistic for θ.

(b) Find a function $\phi = \phi(\theta)$ of θ such that there is an unbiased estimator $\hat{\phi}$ of ϕ with variance $Var(\hat{\phi})$ achieving the Cramer-Rao lower bound. Obtain $\hat{\phi}$.

(c) Is the estimator $\hat{\phi}$ obtained in (b) above the UMVUE (Uniformly Minimum Varainced Unbiased estimator) of ϕ? Why?

10. Consider the following regression model:

$$Y_j = \theta_1 + \theta_2(x_j - \bar{x}) + \epsilon_j, j = 1, \dots, n$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Assume that the ϵ_j's are independently distributed as normal random variables with means 0 and variance σ^2 and that the x_i's are non-stochastic.

(a) Assuming a non-informative prior for $\{\theta_i, i = 1, 2, \sigma^2)$ as $P(\theta_i, i = 1, 2, \sigma^2) \propto (\sigma^2)^{-1}$, derive the posterior distribution of $\{\theta_i, i = 1, 2\}$.

(b) Derive the posterior distribution of θ_2 and a $(1 - \alpha)$ % HPD (Highest Posterior Density) interval for θ_2. How is this HPD interval comparing with the $(1 - \alpha)$ % confidence interval for θ_2 ?

11. Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with density

$$f(x|\mu, \sigma) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, x \geq \mu, -\infty < \mu < \infty, \sigma > 0.$$

(a) Find the maximum likelihood estimates of μ and σ.

(b) Find the method of moment estimates of μ and σ.

12. Let $\{X_1, \ldots, X_n\}$ be a random sample from the normal distribution with mean μ_1 and variance $4\sigma^2$ and $\{Y_1, \ldots, Y_m\}$ a random sample from the normal distribution with mean μ_2 and variance $9\sigma^2$, where σ^2 is unknown.

    (a) Derive a $100(1-\alpha)\%$ confidence interval for $\delta = \mu_1 - \mu_2$.

    (b) Assuming non-informative prior $P(\mu_1, \mu_2, \sigma^2) \propto \sigma^{-2}$, derive a $(1-\alpha)$ % HPD (Highest Posterior Density) Bayesian interval for $\delta = \mu_1 - \mu_2$. How is this HPD interval comparing with the confidence interval obtained in (a) above?

    (c) Illustrate how you will use the above results to test the hypothesis $H_0 : \mu_1 = \mu_2$.

Table of P(Z < z), Z ~N(0,1)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999