PhD Qualifying Exam in Statistics
August, 2007

You need at least 42 points to pass.

1. Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ be two independent random samples. We assume that $X_1, X_2, \dots, X_n$ are independent identically distributed exponential (λ) random variables and $Y_1, Y_2, \dots, Y_m$ are independent identically distributed exponential (μ) random variables. We want to test $H_0: \ \lambda = \mu$ against the alternative $H_A: \ \lambda \neq \mu$.
(i) Find $T_{m,n}$ the likelihood ratio test. (7 points)
(ii) Prove that $-2 \log T_{m,n}$ converges in distribution under the null hypothesis when $\min(m,n) \to \infty$. (7 points)
(iii) Show that the likelihood ratio test is asymptotically consistent and prove that the power goes to 1 as $\min(m,n) \to \infty$. (7 points)

2. Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ be two independent random samples. We assume that $X_1, X_2, \dots, X_n$ are independent identically distributed normal (μ_1, σ^2) random variables and $Y_1, Y_2, \dots, Y_m$ are independent identically distributed normal (μ_2, σ^2) random variables (σ^2 is unknown). We want to test $H_0: \ \mu_1 = \mu_2$ against the alternative $H_A: \ \mu_1 \neq \mu_2$.
Show that the likelihood ratio test and the two sample t–test are equivalent. (7 points)

3. Let $X_1, X_2, \dots, X_n$ be independent identically distributed random variables with distribution function F. We use

$$\hat{f}_n(t) = \frac{1}{nh} \sum_{1 \le i \le n} K\left(\frac{t - X_i}{h}\right)$$

to estimate $f = F'$.

We assume that the following conditions are satisfied:

$$\int_{-1}^{1} K(u)du = 1, \quad K(t) = 0, \quad \text{if} \ \ t \notin [-1, 1]$$

F is three times differentiable and the derivatives are uniformly bounded on the real line

$$h = h(n) \to 0$$

(i) Compute the bias of $\hat{f}_n(t)$ for any t and any n. (7 points)
(ii) Prove that $\hat{f}_n(t)$ is asymptotically unbiased for any t as $n \to \infty$. (7 points)
(iii) Compute the mean square error of $\hat{f}_n(t)$ (7 points)

4. Let X be a random variable with distribution function F. Prove that X is symmetric around 0 (i.e. $F(x) = 1 - F(-x)$ for all x) if and only if the characteristic function of X is real. (7 points)

5. Let X_1 and X_2 be independent standard normal random variables. Find the necessary and sufficient condition for the independence of $Y_1 = a_1X_1 + a_2X_2$ and $Y_2 = b_1X_1 + b_2X_2$. (7 points)

6. Let Let $X_1, X_2, \ldots, X_n$ be independet identically distributed random variables with distribution function
$$F(t) = \left(\frac{t}{\theta}\right)^2 \quad \text{if } 0 \le t \le \theta.$$
(i) Find the maximum likelihood estimator for θ. (7 points)
(ii) Compute the limit distribution of the maximum likelihood estimator. (7 points)

7. Let Let $X_1, X_2, \ldots, X_n$ be independet identically distributed random variables with $EX_1 = \mu$, $\mathrm{var}X_1 = \sigma^2$. We assume that $EX_4 < \infty$. The random variable
$$\hat{\sigma}_n^2 = \frac{1}{2n} \sum_{2 \le i \le n} (X_i - X_{i-1})^2$$
is suggested to estimate σ^2.
(i) Establish some basic properties of $\hat{\sigma}_n^2$. (7 points)
(ii) Is $\hat{\sigma}_n^2$ a better estimator than the sample variance? Justify your answer. (7 points)

8. Let $(X_1, \ldots, X_m)$ be a multinomial random vector with parameters N and $p_1, p_2, \ldots, p_m$. Note that $N = X_1 + \ldots + X_m$ and $p_1 + p_2 + \ldots + p_m = 1$. Find the maximum likelihood estimators for $p_1, p_2, \ldots, p_m$. (7 points)

9. Let X and Y be independent random variables with density function f and g. Assume that $Y > 0$. Show that the function
$$h(t) = \int_0^\infty z f(zt) g(z) dz$$
can be chosen as the density of $Z = X/Y$. (7 points)