REAL ANALYSIS
(Dr. Mieczyslaw K. Dabkowski)
Qualifying Exam
April 11, 2011

Name_____

Instructions. *Please solve any* **five problems** *from the list of the following problems (show all your work).*

1. Show that the following family of subsets of $\mathbb{R}$:

$$\mathcal{M} = \{E \subset \mathbb{R} \mid E \text{ is countable or } E^c \text{ is countable}\}$$

is a $\sigma-$ algebra in $\mathbb{R}$.

2. Recall, the following definition of measurable function:

   Definition *Let M is a σ-algebra in X. A function $h : X \to [-\infty, \infty]$ is measurable if the set $\{x \in X \mid h(x) \geq r\}$ is measurable for every $r \in \mathbb{R}$.*

   Suppose that $f, g : \mathbb{R} \to [-\infty, \infty]$ are measurable functions. Using the definition stated above, show that the following set
   $$\{x \in \mathbb{R} \mid f(x) < g(x)\}$$

   is measurable.

3. Suppose that $f_n : \mathbb{R} \to [0, \infty]$ is measurable for $n = 1, 2, 3, ...$, and $f_1 \geq f_2 \geq ... \geq 0$, $f_n(x) \to f(x)$ as $n \to \infty$, for every $x \in \mathbb{R}$, and $f_1 \in L^1(\mu)$, where μ is the Lebesgue measure. Prove that then

$$\lim_{n \to \infty} \int_{\mathbb{R}} f_n \, d\mu = \int_{\mathbb{R}} f \, d\mu$$

and show that this conclusion does not follow if the condition "$f_1 \in L^1(\mu)$" is omitted.

4. Suppose $\mu(X) = 1$ and suppose f and g are positive measurable functions on X such that $fg \geq 1$. Prove that

$$\left(\int_X f \, d\mu \right) \left(\int_X g \, d\mu \right) \geq 1.$$

5. Let μ is the Lebesgue measure on $\mathbb{R}$. Suppose $f \in L^1(\mu)$. Prove that to each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\int_E |f| \, d\mu < \epsilon$$

whenever $\mu(E) < \delta$.

6. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a measurable function on $\mathbb{R}$ and let μ be the Lebesgue measure on $\mathbb{R}$. Define the following function

$$\varphi(p) = \int_{\mathbb{R}} |f|^p \, d\mu = \| f \|_p^p, \quad (0 < p < \infty).$$

Let $E = \{p \mid \varphi(p) < \infty\}$ and assume that $\| f \|_\infty > 0$. If $r < p < s$, $r \in E$, and $s \in E$, prove that $p \in E$.

Ph.D. Qualifying Examination in Probability

April 11, 2011

Instructions:

(a) *There are three problems, each of equal weight. You may submit work on all three.*

(b) *Extra credit will be given for a problem with all parts solved well.*

(c) *Look over all three problems before beginning work.*

(d) *Start each problem on a new page, and number the pages.*

(e) *On each page, indicate problem number and part, and write your name.*

(f) *Indicate your lines of reasoning and what background results are being applied.*

---

1. Let $X_1, X_2, \ldots$ be independent random variables with means $\mu_1, \mu_2, \ldots$ and finite variances $\sigma_1^2, \sigma_2^2, \ldots$ not necessarily bounded. Suppose that $\sum_{i=1}^{\infty} \sigma_i^2/i^2 < \infty$. Exhibit, for any $\varepsilon > 0$, an upper bound to the probability

$$P\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} \sum_{i=1}^{n} \mu_i \right| > \varepsilon \right)$$

that converges to 0 as $n \to \infty$.

---

2. Let $S_n = X_1 + \ldots + X_n$ be a sequence of sums of non-negative random variables X_n, and suppose that S_n converges to a random variable S in probability.

(a) Can you conclude that S_n converges to S almost surely?

(b) If $E(S^p) < \infty$ for some $p > 0$, can you conclude that

   – S_n converges to S in L^p (in mean of order p)?
   – $E(S_n^p)$ converges to $E(S^p)$?

Justify your answers.

---

3. Two gamblers play the following game. Player I tosses a fair coin. If it lands up heads, he pays \$1 to player II. Otherwise, he receives \$1 from player II. Player I starts with \$$n$, player II starts with \$$m$, and the game continues until one of the players runs out of money.

(a) Compute the probability that player I wins (i.e., player II runs out of money).

(b) Compute the expected duration of the game.

Hints. For (a), consider sequence Y_n, which is the profit (positive or negative) that player I has after n coin tosses and show that it is a martingale. For (b), consider sequence $(Y_n^2 - n)$ and show that it is a martingale.

Ph.D. Qualifying Examination in Statistical Inference

April 15, 2011

Instructions:

(a) *There are three problems, each of equal weight. You may submit work on all three.*

(b) *Extra credit will be given for a problem with all parts solved well.*

(c) *Look over all three problems before beginning work.*

(d) *Start each problem on a new page, and number the pages.*

(e) *On each page, indicate problem number and part, and write your name.*

(f) *Indicate your lines of reasoning and what background results are being applied.*

1. Let $\pi(\theta)$ be a probability density function for parameter θ supported on Ω, and let $R(\theta, \delta)$ denote the risk function (expected loss, given θ) of any decision procedure δ using data X. Suppose that the Bayes rule $\delta_\pi(X)$ corresponding to $\pi(\theta)$ satisfies

$$\int_\Omega R(\theta, \delta_\pi)\pi(\theta)d\theta = \sup_{\theta \in \Omega} R(\theta, \delta_\pi).$$

Explaining all steps and stating all definitions used, prove or disprove each of the following assertions:

(i) δ_π is minimax.

(ii) If δ_π is the unique Bayes rule with respect to $\pi(\theta)$, then δ_π is uniquely minimax.

(iii) $\pi(\theta)$ is the least favorable probability density function for θ.

2. Let us have a sample $X_1, \ldots, X_n$ from a population with probability density function

$$f(x|\theta) = (2\theta)^{-1}I(|x| < \theta), \quad \theta > 0.$$

(i) Find a minimum variance unbiased estimator of θ.

(ii) Explore uniqueness of this estimator.

3. Let us have a sample $X_1, \ldots, X_n$ from the distribution Normal$(\theta, A\theta)$, where both A and θ are unknown. Find the $1 - \alpha$ confidence set for the parameter A that is obtained by inverting the likelihood ratio test of $H_0 : A = A_0$ versus $H_a : A \neq A_0$.

- Number of questions = 3. Answer all of them. Total points = 50.

- Simplify your answers as much as possible and carefully justify all steps to get full credit.

- There is no need to prove any standard result. Just state the result and use it.

- **All vectors are column vectors.**

1. Consider the linear model

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where $E(\epsilon_i) = 0$, $\mathrm{var}(\epsilon_i) = \sigma^2 x_i$, the errors are uncorrelated, and all $x_i > 0$. Note that Y_i, x_i, ϵ_i and β in this problem are all scalar quantities.

   (a) [8 points] Show that the weighted least squares estimator of β is $\hat{\beta} = \overline{Y}/\overline{x}$.

   (b) [7 points] Show that $\mathrm{var}(\hat{\beta}) = \sigma^2 / \sum x_i$.

2. Consider the following one-way ANOVA model with three groups and n observations per group:
$$Y_{ij} = \theta_i + \epsilon_{ij}, \quad j = 1, \ldots, n, \quad i = 1, 2, 3.$$

Here we assume that the errors are independently and identically distributed as $N(0, \sigma^2)$ random variables, and σ^2 is _known_. Let $\overline{Y}_i$ denote the sample mean of ith group.

   (a) [7 points] Find the joint distribution of the vector $(\overline{Y}_2 - \overline{Y}_1, \overline{Y}_3 - \overline{Y}_1)$. Be sure to name the distribution and specify the parameters of the distribution.

   (b) [8 points] Suppose c_α represents the $(1 - \alpha)$th percentile of the scalar random variable $\max\{Z_1, Z_2\}$, where the random vector (Z_1, Z_2) follows a bivariate normal distribution with standard normal marginals and correlation $1/2$. Define the random variables

$$L_{21} = \overline{Y}_2 - \overline{Y}_1 - c_\alpha \sigma \sqrt{2/n}, \quad L_{31} = \overline{Y}_3 - \overline{Y}_1 - c_\alpha \sigma \sqrt{2/n}.$$

Use the distribution derived in (a) to show that L_{21} and L_{31} represent $100(1-\alpha)\%$ simultaneous _lower_ confidence bounds for $\theta_2 - \theta_1$ and $\theta_3 - \theta_1$, respectively. In other words, show that

$$P(\theta_2 - \theta_1 \geq L_{21}, \ \theta_3 - \theta_1 \geq L_{31}) = 1 - \alpha.$$

3. Suppose the vector (X_1, X_2) follows a bivariate normal distribution with mean (μ_1, μ_2), variance (σ_1^2, σ_2^2) and covariance σ_{12}. Let $D = X_1 - X_2$ and $S = X_1 + X_2$. We want to derive a test of the *joint* null hypothesis $H_0 : \mu_1 = \mu_2, \sigma_1^2 = \sigma_2^2$, by regressing D on S.

(a) [5 points] Show that the joint distribution of the vector (D, S) is bivariate normal. Express the parameters of (D, S) in terms of parameters of (X_1, X_2).

(b) [5 points] Show that $E(D|S) = \beta_0 + \beta_1 S$, where

$$\beta_0 = (\mu_1 - \mu_2) - \left\{ \frac{\sigma_1^2 - \sigma_2^2}{\text{var}(S)} \right\} (\mu_1 + \mu_2) \quad \text{and} \quad \beta_1 = \frac{\sigma_1^2 - \sigma_2^2}{\text{var}(S)}.$$

(c) [3 points] Use the result in (b) to find a null hypothesis that is equivalent to the desired H_0.

(d) [7 points] Suppose we have n independent observations of (X_1, X_2), namely, (X_{1i}, X_{2i}), $i = 1, \ldots, n$. Use these data and the result in (c) to derive a test for the desired H_0. Be sure to specify the test statistic, its null distribution along with the degrees of freedom, and the rejection region of the test.

Ph.D. Qualifying Examination in Statistical Methods

April 18, 2011

A real estate appraiser is interested in predicting residential home prices in a mid-western city as a function of various features.

Data on 522 recent home sales are available on the enclosed CD and also on the web site http://www.utdallas.edu/~mbaron/Qual/. All the three files contain identical data in different formats. The following variables are included.

Column	Variable
1	Identification number 1–522
2	Sales price of residence (dollars)
3	Finished area of residence (square feet)
4	Total number of bedrooms in residence
5	Total number of bathrooms in residence
6	Air conditioning: present or absent
7	Number of cars that garage will hold
8	Pool: present or absent
9	Year property was originally constructed
10	Quality of construction: high, medium, or low
11	Indicator of architectural style
12	Lot size (square feet)
13	Location near a highway: yes or no

- Develop the best model you can for predicting the home sales prices. Use the suitable variable selection and regression diagnostics methods.

- If any of your conclusions are based on certain assumptions, state them and verify their validity. Apply remedial measures if necessary.

- Test whether any interaction exists between the construction quality, air conditioning, and the presence of a pool.

- As a separate task, derive the best model you can for predicting the construction quality of a home.

Instructions

– Load the data and conduct the necessary data analysis using *software of your choice.*

– Submit a report, written or typed, hard copy or e-mail. If you choose to e-mail the report, send it to both ammann@utdallas.edu and mbaron@utdallas.edu.

– In the report, describe every step of your analysis: methods, reasons, results, and conclusions. For example:

Test significance of variable .... Use SAS, PROC ... with option ... The F test gives a p-value .... Therefore, ...

Verify assumptions of the test. Use ... ... Variable ... violates assumption ... because ... Therefore, ...

– Attach your computer programs and *only relevant* parts of the output.