Qualify Exam (June 10, 2014)

STAT5005+STAT5010 (09:00am-12:00pm)

STAT5005 Advanced Probability Theory

1.
   a) Prove that $W_n \xrightarrow{d} Z$ iff $\mathbb{E}f(Wn) \to \mathbb{E}f(Z)$ for any bounded continuous function f.
   b) $X_1, X_2, \ldots$ are iid with mean 0 and variance 1. Prove that

$$\lim_{n\to\infty} \frac{1}{n^{1/2}} \mathbb{E}\left[\sum_{i=1}^{n} |X_i|\right] = \sqrt{\frac{2}{\pi}}.$$

2.
   a) Prove the Marcinkiewicz-Zygmund strong law of large numbers ($\mathbb{E}(\mathrm{X}) = 0, 1 < p < 2$).
   b) $X_1, X_2, \ldots$ are iid with mean 0 and variance 1. $S_n = X_1 + \cdots + X_n$. Prove that

$$\operatorname*{limsup}_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = 1, a.s..$$

3.
   a) Prove that if A_1 and A_2 are independent, then $\sigma(A_1)$ and $\sigma(A_2)$ are independent.
   b) $X_1, X_2, \ldots$ are iid with $\mathbb{P}(X_i \le x) = e^{-x}$. $S_n = X_1 + \cdots + X_n$. Find the limiting distribution of $\sum_{i=1}^{n} I(X_i S_n > 1)$.
   c) $X_1, X_2, \ldots$ are iid Unif(0,1) random variables. $S_n = X_1 + \cdots + X_n$. Let $T = \inf\{n: S_n > 1\}$. Find $\mathbb{P}(T > n), \mathbb{E}(T), \mathbb{E}(S_T)$.

STAT5010 Advanced Statistical Inference

1. $X_1, \ldots, X_n$ sample from $N(\theta, \sigma^2)$
   a) If $\sigma^2 = \sigma_0^2$ known, prove $\bar{X}$ is UMVUE of θ
   b) If σ^2 is unknown, prove $\bar{X}$ is still UMVUE of θ by noting that $\bar{X}$ doesn't depend on σ_0^2.
   c) If $\theta = \theta_0$ known, show that $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is not UMVUE of σ^2.
   d) If $\sigma^2 = \sigma_0^2$ known, find the UMVUE of $\mathbb{P}(\mathrm{X}_1 \ge 0)$.
2. Two-sample test with equal variance. Derive LRT for $H_0: \mu_X - \mu_Y \ge \delta \leftrightarrow H_1: \mu_X - \mu_Y < \delta$.

STAT5020+STAT5030 (02:00pm-04:00pm)

STAT5020 Topics in Multivariate Analysis

1. Analysis of the multivariate heterogeneous data.
   a) Propose an appropriate model.
   b) Describe the Bayesian analysis of the proposed model.
   c) State the model comparison in the content of Bayes factor.
2.
   a) How many types of missingness? Which one is ignorable? Which one is nonignorable? Why?
   b) Describe how to analyze the longitudinal data in the presence of nonignorable missing data.

STAT5030 Linear Models

1. $y_i = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_{p-1} x_{p-1,i} + \epsilon_i, i = 1, \dots, n$

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

   Assume $\beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$.
   a) Find the distribution of R^2.
   b) Find the value of $\mathbb{E}[R^2]$.
   c) Find the value of $\mathrm{Var}[R^2]$.
2. $Y = X\beta + \epsilon, \epsilon \sim N(0, \sigma^2 V)$. $x_i > 0$.

$$V = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & 1 \end{pmatrix}$$

   Let $\hat{\beta}$ be OLS estimator of β.
   a) If $\rho = 0$. Define $G_1 = Var(\hat{\beta})$. Find G_1.
   b) If $\rho > 0$. Define $G_2 = Var(\hat{\beta})$. Find G_2.
   c) Which one of G_1 and G_2 is larger? Why?
   d) Let k_1 and k_2 are two constant vectors. Discuss how to construct $100(1-\alpha)$-CI of

$$\frac{k_1' \beta}{k_2' \beta}.$$