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

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = 1, a.s..$$

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

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

2014 Qualify Reference Answer

Chaojie Wang

June 8, 2015

1 5005

1.(a) Theorem 3.2.3. If $W_n \Rightarrow Z$, then there exist $Y_n =_d W_n$ so that $Y_n \to Y$ a.s. For any bounded continuous function f, $f(Y_n) \to f(Y)$ a.s.. Using bounded convergence theorem, then $Ef(Y_n) \to Ef(Y)$. Thus, $Ef(W_n) = Ef(Y_n) \to Ef(Y) = Ef(Z)$.

If for any bounded continuous function f, $Ef(W_n) \to Ef(Z)$. Let

$$f_{x,\epsilon}(y) = \begin{cases} 1, & y \le x \\ 1 - (y-x)/\epsilon, & x \le y \le x+\epsilon \\ 0, & y \ge x+\epsilon \end{cases}$$

here $f_{x,\epsilon}$ is continuous, so $0 \le f_{x,\epsilon}(y) \le 1$ on $y \in [x, x+\epsilon]$, thus

$$\limsup_{n\to\infty} P(W_n \le x) \le \limsup_{n\to\infty} Ef_{x,\epsilon}(W_n) = Ef_{x,\epsilon}(Z) \le P(Z \le x+\epsilon)$$

Let $\epsilon \downarrow 0$, then $\limsup_{n\to\infty} P(W_n \le x) \le P(Z \le x)$. On the other hand,

$$\liminf_{n\to\infty} P(W_n \le x+\epsilon) \ge \liminf_{n\to\infty} Ef_{x,\epsilon}(W_n) = Ef_{x,\epsilon}(Z) \ge P(Z \le x)$$

Let $\epsilon \downarrow 0$, then $\liminf_{n\to\infty} P(W_n \le x) \ge P(Z \le x)$. Thus, $P(W_n \le x) \to P(Z \le x)$. It implies $W_n \Rightarrow Z$.

(b) The correct question should be to prove

$$\lim_{n\to\infty} \frac{1}{n^{1/2}} E|\sum_{i=1}^n X_i| = \sqrt{\frac{2}{\pi}}$$

Let $Y_n = |\sum_{i=1}^n X_i / n^{1/2}|$. It suffices to prove $EY_n \to E|Z|$ where Z is standard normal distribution. Fixed any $R > 0$,

$$|EY_n - E|Z|| = |\int_0^\infty P(Y_n > t)dt - \int_0^\infty P(|Z| > t)dt|$$
$$\leq |\int_0^R P(Y_n > t) - P(|Z| > t)dt| + |\int_R^\infty P(Y_n > t)dt| + |\int_R^\infty P(|Z| > t)dt|$$
$$\leq \int_0^R |P(Y_n > t) - P(|Z| > t)|dt + \int_R^\infty P(Y_n > t)dt + \int_R^\infty P(|Z| > t)dt$$

Since $EY_n^2 = 1$, using Cauchy-Schwarz inequality,

$$\int_R^\infty P(Y_n > t)dt = \int_R^\infty P(Y_n 1_{(Y_n > R)} > t)dt \leq E(Y_n 1_{(Y_n > R)})$$
$$\leq (EY_n^2 E1^2_{(Y_n > R)})^{1/2} = (P(Y_n > R))^{1/2}$$

Using Chebyshev's inequality,

$$(P(Y_n > R))^{1/2} \leq (\frac{1}{R^2} EY_n^2)^{1/2} = \frac{1}{R}$$

Thus, $\int_R^\infty P(Y_n > t)dt \leq \frac{1}{R}$. Similarly, $\int_R^\infty P(|Z| > t)dt \leq \frac{1}{R}$. Thus,

$$|EY_n - E|Z|| \leq R \sup_{t \in [0,R]} |P(Y_n > t) - P(|Z| > t)| + \frac{2}{R}$$

By central limit theorem, we have $\sum_{i=1}^n X_i/n^{1/2} \Rightarrow Z$. Since $f(x) = |x|$ is continuous function only except the point $x = 0$. Thus, $|\sum_{i=1}^n X_i/n^{1/2}| \Rightarrow |Z|$, it implies that

$$\sup_{t \in [0,R]} |P(Y_n > t) - P(|Z| > t)| \to 0 \ as \ n \to \infty$$

Let R go to infinity, then $EY_n \to E|Z| = \sqrt{2/\pi}$. Proven.

Remark: Since $f(x) = |x|$ is not bounded function, the expectation and limit operation could not be exchanged directly.

2.(a) Theorem 2.5.8. Let $Y_k = X_k 1_{(|X_k| \leq k^{1/p})}$. Since $E|X_1|^p < \infty$, then

$$\sum_{k=1}^\infty P(X_k \neq Y_k) = \sum_{k=1}^\infty P(|X_k| > k^{1/p}) \leq \int_0^\infty P(|X_k| > x^{1/p})dx$$
$$= \int_0^\infty P(|X_k|^p > x)dx = E|X_1|^p < \infty$$

Using Borel-Cantelli lemma, $P(X_n \neq Y_n \ i.o.) = 0$. Thus, let $T_n = Y_1 + \cdots + Y_n$, it suffices to prove $T_n/n^{1/p} \to 0$ a.s. Since $P(|Y_m| > y) \leq P(|X_1| > y)$,

$$\sum_{m=1}^\infty Var(Y_m/m^{1/p}) \leq \sum_{m=1}^\infty EY_m^2/m^{2/p} = \sum_{m=1}^\infty \int_0^\infty \frac{1}{m^{2/p}} 2yP(|Y_m| > y)dy$$
$$= \sum_{m=1}^\infty \sum_{n=1}^m \int_{(n-1)^{1/p}}^{n^{1/p}} \frac{1}{m^{2/p}} 2yP(|Y_m| > y)dy \leq \sum_{m=1}^\infty \sum_{n=1}^m \int_{(n-1)^{1/p}}^{n^{1/p}} \frac{1}{m^{2/p}} 2yP(|X_1| > y)dy$$
$$= \sum_{n=1}^\infty \int_{(n-1)^{1/p}}^{n^{1/p}} \{\sum_{m=n}^\infty \frac{1}{m^{2/p}}\} 2yP(|X_1| > y)dy$$

Since $1 < p < 2$, then $x^{1-2/p} \to 0$ as $x \to \infty$, when $y \in ((n-1)^{1/p}, n^{1/p})$, then

$$\sum_{m=n}^{\infty} \frac{1}{m^{2/p}} \leq \int_{n-1}^{\infty} x^{-2/p} dx = \frac{p}{p-2}(n-1)^{1-2/p} \leq C y^{p-2}$$

Thus,

$$\sum_{m=1}^{\infty} Var(Y_m/m^{1/p}) \leq \sum_{n=1}^{\infty} \int_{(n-1)^{1/p}}^{n^{1/p}} C y^{p-2} 2y P(|X_1| > y) dy$$
$$= \int_0^{\infty} 2C y^{p-1} P(|X_1| > y) dy \leq C' E|X_1|^p < \infty$$

It implies that $\sum_{m=1}^{\infty} Y_m/m^{1/p}$ converge a.s. Since $m^{1/p} \uparrow \infty$, then

$$n^{-1/p} \sum_{m=1}^{n} (Y_m - \mu_m) \to 0 \ a.s.$$

Note that $\mu_m = EY_m$, since $EX_m = 0$, then $\mu_m = -E(X_m; |X_m| > m^{1/p})$, so

$$|\mu_m| \leq E(|X_m|; |X_m| > m^{1/p}) = m^{1/p} E(|X_m|/m^{1/p}; |X_m| > m^{1/p})$$
$$\leq m^{1/p} E(|X_m|^p/m; |X_m| > m^{1/p}) = m^{1/p-1} E(|X_m|^p; |X_m| > m^{1/p})$$

Now $\sum_{m=1}^{n} m^{1/p-1} \leq C n^{1/p}$ and $E(|X_m|^p; |X_m| > m^{1/p}) \to 0$ as $m \to \infty$,

Thus, $n^{-1/p} \sum_{m=1}^{n} \mu_m \to 0$. It implies $n^{-1/p} \sum_{m=1}^{n} Y_m \to 0$ a.s. Proven.

(b) This problem is called Law of iterated logarithm. Its proof is much more difficult and I think it could not be finished in limited time. Here just list some outline of proof.

It suffices to prove, for any $\epsilon > 0$,

$$P(S_n > (1+\epsilon)\sqrt{2n \log\log n} \ i.o.) = 0$$
$$P(S_n > (1-\epsilon)\sqrt{2n \log\log n} \ i.o.) = 1$$

Using central limit theorem, $S_n/\sqrt{n} \Rightarrow N(0,1)$. Since theorem 1.2.3,

$$\int_x^{\infty} e^{-y^2/2} dy \sim \frac{1}{x} e^{-x^2/2}$$

Thus, for the upper bound,

$$P(S_n/\sqrt{n} > (1+\epsilon)\sqrt{2 \log\log n}) \leq C e^{-(1+\epsilon)^2 2 \log\log n/2} = C(\log n)^{-(1+\epsilon)^2}$$

Using the estimate of n of the form $a^k, a > 1$, then the probability is summable. Then using Borel-Cantelli lemm, the subsequence is proven. For the lower bounded is similarly.

3.(a) Another condition is $\mathcal{A}_1$ and $\mathcal{A}_2$ are π-system. Theorem 2.1.3.

For any $A \in \mathcal{A}_1$, Construct $\mathcal{B} = \{B : P(A)P(B) = P(A \cap B)\}$. Since $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent, obviously that for any $B \in \mathcal{A}_2$, then $B \in \mathcal{B}$. Thus, $\mathcal{A}_2 \subset \mathcal{B}$.

Next, check $\mathcal{B}$ is λ-system. (i) $P(A)P(\Omega) = P(A) = P(A \cap \Omega)$ implies $\Omega \in \mathcal{B}$.

(ii) If $B \in \mathcal{B}$, then $P(A \cap B) = P(A)P(B)$. Thus,

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A)(1 - P(B)) = P(A)P(B^c)$$

So, $B^c \in \mathcal{B}$.

(iii) If $B_i \in \mathcal{B}$ are disjoint countable sequence, then $B_i \cap A$ are disjoint, so,

$$P((\cup_i B_i) \cap A) = P(\cup_i (B_i \cap A)) = \sum_i P(B_i \cap A) = \sum_i P(B_i)P(A) = P(\cup_i B_i)P(A)$$

It implies that $\cup_i B_i \in \mathcal{B}$.

Thus, $\mathcal{B}$ is λ-system containing $\mathcal{A}_\in$ and $\mathcal{A}_2$ is π-system, using π-λ theorem, $\sigma(\mathcal{A}_2) \subset \mathcal{B}$. Thus, for any $B \in \sigma(\mathcal{A}_2)$, then $P(A \cap B) = P(A)P(B)$ for any $A \in \mathcal{A}_1$, it implies $\mathcal{A}_1$ and $\sigma(\mathcal{A}_2)$ are independent.

Repeat this process, it implies that $\sigma(\mathcal{A}_1)$ and $\sigma(\mathcal{A}_2)$ are independent.

(b). Here the problem should be to find the limiting distribution of $\sum_{i=1}^n I(X_i S_n < 1)$.

And this problem is still confused since each $I_{(X_i S_n)<1}$ may be dependent. Here we view them independent.

Let $Y_i = I(X_i S_n < 1)$ with $P(Y_i = 1) = P(X_i S_n < 1) = p_i$ and $P(Y_i = 0) = 1 - p_i$. Since Y_i are not independent, consider Poisson convergence, it suffice to prove

(i) $\sum_{i=1}^n p_i \to \lambda$ and (ii) $\max_i p_i \to 0$.

For (i), let $S_n^i = S_n - X_i$, then X_i and S_n^i are independent, we have

$$\begin{aligned} p_i &= EI(X_i S_n < 1) = E(E(I(X_i S_n < 1))|S_n^i) = E(P(X_i S_n < 1)|S_n^i) \\ &= E(P(X_i^2 + X_i S_n^i < 1)|S_n^i) = E(P(0 \le X_i \le g(S_n^i))|S_n^i) \end{aligned}$$

where $g(S_n^i) = \frac{-S_n^i + \sqrt{{S_n^i}^2 + 4}}{2} = \frac{2}{S_n^i + \sqrt{{S_n^i}^2 + 4}}$. Since $S_n^i \to n$ by law of large numbers, then $g(S_n^i) \sim 2/n$. Since X_i follow exponential(1), then

$$p_i = 1 - exp(-g(S_n^i)) \approx g(S_n^i) \sim 2/n$$

Thus, $\sum_{i=1}^n p_i \to 2$.

For (ii), since for each i, p_i are equivalent, then $\max_i p_i \sim 2/n \to 0$

Thus, $S_n \sim Poisson(2)$

(c). Since $X_i \in [0, 1]$, then $\{T = n\} = \{S_{n-1} \le 1, S_n > 1\}$. Then $\{T = n\}$ for all n are disjoint. Thus

$$P(T > n) = \sum_{k=n+1}^{\infty} P(T = k) = \sum_{k=n+1}^{\infty} P(S_{k-1} \le 1, S_k > 1) = P(S_n \le 1)$$

So

$$
\begin{aligned}
P(S_n \leq 1) &= \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-x_1\cdots-x_{n-1}} dx_n \cdots dx_2 dx_1 \\
&= \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-x_1\cdots-x_{n-2}} (1-x_1\cdots-x_{n-1}) dx_{n-1} \cdots dx_2 dx_1 \\
&= \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-x_1\cdots-x_{n-3}} \frac{1}{2}(1-x_1\cdots-x_{n-2})^2 dx_{n-2} \cdots dx_2 dx_1 \\
&= \cdots = \int_0^1 \frac{1}{(n-1)!}(1-x_1)^{n-1} dx_1 = \frac{1}{n!} = P(T > n)
\end{aligned}
$$

And

$$
ET = \sum_{n=0}^{\infty} P(T > n) = \sum_{n=0}^{\infty} \frac{1}{n!} = e
$$

Since $\{T = n\} = \{S_1 < 1, \cdots, S_{n-1} < 1, S_n > 1\} \in \mathcal{F}_n$, then $\{T = n\}$ is stopping time, and $ET = e < \infty$, using Wald's Equation, $ES_T = EX_1 ET = e/2$

2 5010

1.(a) If $\sigma^2 = \sigma_0^2$ is known,

$$
\begin{aligned}
f(\mathbf{x}|\mu) &= (\frac{1}{2\pi\sigma_0^2})^{n/2} exp\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\theta)^2\} \\
&= (\frac{1}{2\pi\sigma_0^2})^{n/2} exp\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n} x_i^2\} exp\{-\frac{n\theta^2}{2\sigma_0^2}\} exp\{-\frac{n\bar{x}}{\sigma_0^2}\theta\}
\end{aligned}
$$

is exponential family. Thus, $\bar{X}$ is complete and sufficient statistic. And $E\bar{X} = \theta$ is unbiased, thus $\bar{X}$ is UMVUE of θ.

(b) If σ^2 is unknown,

$$
\begin{aligned}
f(\mathbf{x}|\mu) &= (\frac{1}{2\pi\sigma^2})^{n/2} exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\} \\
&= (\frac{1}{2\pi\sigma^2})^{n/2} exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \frac{n}{2\sigma^2}(\bar{x}-\theta)^2\}
\end{aligned}
$$

is exponential family, then $(\bar{X}, S^2)$ is complete and sufficient statistic for (θ, σ^2) where $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2$. And $E\bar{X} = \theta$ is unbiased, thus $E(\bar{X}|\bar{X}, S^2)$ is UMVUE of θ.

It suffices to prove $\bar{X}$ and S^2 are independent. then $E(\bar{X}|\bar{X}, S^2) = \bar{X}$.

Since $\bar{X} = \frac{1}{n}\mathbf{1}'X$ and $S^2 = X'X - X'(\frac{1}{n}\mathbf{1}\mathbf{1}')X = X'(I - \frac{1}{n}J)X$ is quadratic form, and $\frac{1}{n}\mathbf{1}'\Sigma(I - \frac{1}{n}J) = 0$ where $\Sigma = \sigma^2 I$, then $\bar{X}$ and S^2 are independent.

(c) To prove $\frac{1}{n}\sum_{i=1}^n (X_i - \theta)^2$ is unbiased estimator of σ^2 and has smaller variance.

For unbiased estimator,

$$E(\frac{1}{n}\sum_{i=1}^{n}(X_i-\theta)^2) = E(X_i-\theta)^2 = \sigma^2$$

For variance, let $Y = X - \theta \sim N(0, \sigma^2)$

$$Var(\frac{1}{n}\sum_{i=1}^{n}(X_i-\theta)^2) = \frac{1}{n}Var(X_i-\theta)^2 = \frac{1}{n}(EY^4 - (EY^2)^2) = \frac{2}{n}\sigma^4$$

The variance of $\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$, let $Z = X - \bar{X} \sim N(0, \frac{n+1}{n}\sigma^2)$

$$\begin{aligned} Var(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2) &= \frac{n}{(n-1)^2}VarZ^2 = \frac{n}{(n-1)^2}(EZ^4 - (EZ^2)^2) \\ &= \frac{2n}{(n-1)^2}\frac{(n+1)^2}{n^2}\sigma^4 \geq \frac{2}{n}\sigma^4 \end{aligned}$$

(d) If $\sigma^2 = \sigma_0^2$ is known, then $I_{(X_1\geq 0)}$ is unbiased for $P(X_1 \geq 0)$. Then $E(I_{(X_1\geq 0)}|\bar{X})$ is UMVUE of $P(X_1 \geq 0)$. To check $E(I_{(X_1\geq 0)}|\bar{X})$.

$$E(I_{(X_1\geq 0)}|\bar{X}) = P(X_1 \geq 0|\bar{X}) = \int_0^\infty P(X_1 = x|\bar{X})dx$$

where

$$\begin{aligned} P(X_1 = x_1|\bar{X}) &= \frac{P(X_1 = x_1, \sum_{i=1}^n X_i = n\bar{X})}{P(\sum_{i=1}^n X_i = n\bar{X})} = \frac{P(X_1 = x_1, \sum_{i=2}^n X_i = n\bar{X} - x_1)}{P(\sum_{i=1}^n X_i = n\bar{X})} \\ &= \frac{\frac{1}{\sqrt{2\pi}\sigma}exp(-\frac{1}{2\sigma^2}(x_1-\theta)^2)\frac{1}{\sqrt{2\pi}\sqrt{n-1}\sigma}exp(-\frac{1}{2(n-1)\sigma^2}(n\bar{X} - x_1 - (n-1)\theta)^2)}{\frac{1}{\sqrt{2\pi n}\sigma}exp(-\frac{1}{2n\sigma^2}(n\bar{X} - n\theta)^2)} \\ &= \frac{1}{\sqrt{2\pi}\sigma}\sqrt{\frac{n}{n-1}}exp(-\frac{n}{2(n-1)\sigma^2}(\bar{X} - x_1)^2) \sim N(\bar{X}, \frac{n-1}{n}\sigma^2) \end{aligned}$$

2. Likelihood ratio test:

$$LRT = \frac{\sup_{\theta\in\Theta_0} f(\mathbf{x}|\mu_X, \sigma^2)f(\mathbf{y}|\mu_Y, \sigma^2)}{\sup_{\theta\in\Theta} f(\mathbf{x}|\mu_X, \sigma^2)f(\mathbf{y}|\mu_Y, \sigma^2)}$$

For the numerator:

$$f(\mathbf{x}|\mu_X, \sigma^2)f(\mathbf{y}|\mu_Y, \sigma^2) = (\frac{1}{\sqrt{2\pi}\sigma})^{m+n}exp(-\frac{1}{2\sigma^2}(\sum_{i=1}^{m}(X_i-\mu_X)^2 + \sum_{i=1}^{n}(Y_i-\mu_Y)^2))$$

Take logarithm and derivative, under $\mu_X - \mu_Y \geq \delta$,

$$\begin{aligned} \hat{\mu}_X &= \frac{m\bar{X} + n\bar{Y} + n\delta}{m+n}, \quad \hat{\mu}_Y = \frac{m\bar{X} + n\bar{Y} - m\delta}{m+n} \\ \hat{\sigma}^2 &= \frac{1}{m+n}(\sum_{i=1}^{m}(X_i-\hat{\mu}_X)^2 + \sum_{i=1}^{n}(Y_i-\hat{\mu}_Y)^2) \end{aligned}$$

For the dominator, take logarithm and derivative,

$$\hat{\mu}_X = \bar{X}, \quad \hat{\mu}_Y = \bar{Y}, \quad \hat{\sigma}_0^2 = \frac{1}{m+n}(\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2)$$

Thus,

$$\begin{aligned}\lambda &= (\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2})^{-\frac{m+n}{2}} = (\frac{\sum_{i=1}^{m}(X_i - \hat{\mu}_X)^2 + \sum_{i=1}^{n}(Y_i - \hat{\mu}_Y)^2}{\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2})^{-\frac{m+n}{2}} \\ &= (1 + \frac{(\bar{X} - \bar{Y} + \delta)^2/(\frac{1}{m} + \frac{1}{n})}{\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2})^{-\frac{m+n}{2}} = (1 + \frac{1}{m+n-2}T^2)^{-\frac{m+n}{2}}\end{aligned}$$

where

$$T = \frac{(\bar{X} - \bar{Y} + \delta)/\sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{(\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2)/(m+n-2)}} \sim t_{m+n-2}$$

The rejection region is $\{(\mathbf{x}, \mathbf{y}) : \lambda(\mathbf{x}, \mathbf{y}) < c\}$ with $P(\lambda(\mathbf{x}, \mathbf{y}) < c) = \alpha$.

3 5020

1. Heterogeneous data means mixture data.
2. MAR and non-ignorable.

4 5030

1. For the numerator, since $H = X(X'X)^{-1}X'$ with column is $\mathbf{1}$, then $HJ = JH = J$,

$$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = (\hat{Y} - 1\frac{1}{n}1'Y)'(\hat{Y} - 1\frac{1}{n}1'Y) = Y'(H - \frac{1}{n}J)(H - \frac{1}{n}J)Y = Y'(H - \frac{1}{n}J)Y$$

For the dominator,

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = (Y - \frac{1}{n}JY)'(Y - \frac{1}{n}JY) = Y'(I - \frac{1}{n}J)(I - \frac{1}{n}J)Y = Y'(I - \frac{1}{n}J)Y$$

Thus,

$$R^2 = \frac{Y'(H - \frac{1}{n}J)Y}{Y'(I - \frac{1}{n}J)Y} = \frac{1}{\frac{Y'(I-\frac{1}{n}J)Y}{Y'(H-\frac{1}{n}J)Y}} = \frac{1}{1 + \frac{Y'(I-H)Y}{Y'(H-\frac{1}{n}J)Y}}$$

Since $(I - H)(H - \frac{1}{n}J) = 0$, then $Y'(I - H)Y$ and $Y'(H - \frac{1}{n}J)Y$ are independent. Then let $F = Y'(I - H)Y/Y'(H - \frac{1}{n}J)Y \sim F_{n-r(X),r(X)-1}$. So $R^2 = 1/(1 + F)$.

It seems there is no closed for of $E(R^2)$ and $Var(R^2)$

2. Here matrix V should be $n \times n$ matrix.

(a) If $\rho = 0$, then $G_1 = Var(\hat{\beta}) = (X'X)^{-1}\sigma^2$

(b) If $\rho > 0$, then $V^{-1/2}Y = V^{-1/2}X\beta + \delta$ where $\delta \sim N(0, \sigma^2 I)$.

Then $G_2 = (X'V^{-1}X)^{-1}(X'V^{-1}VV^{-1}X)(X'V^{-1}X)^{-1}\sigma^2 = (X'V^{-1}X)^{-1}\sigma^2$

(c) G_1 is larger then G_2. Since if the data are dependent, it would generate less random effect. Thus, the variance of estimator would be smaller.

(d) Let $k_1'\beta/k_2'\beta = t$ then $K'\beta = (k_1 - tk_2)'\beta = 0$.

Construct hypothesis test $H_0 : K'\beta = 0$, let $Q = (K'\hat{\beta})'(K'(X'X)^{-1}K)^{-1}K'\hat{\beta}$, we have $F = Q/\hat{\sigma}^2 \sim F_{1,n-r(X)}$. Then reject region is $F > F_{1,n-r(X),\alpha}$.

Construct $100(1-\alpha)\%$ CI, $Q/\hat{\sigma}^2 \leq F_{1,n-r(X),\alpha}$. Then solve the range of t. that is confidence interval.

- MLE
- UMVUE (Important)
  CR lower bound
  L-S theorem
- Testing
  simple test

3 Final(5005)

1. (a) $\mathcal{A}_1$ and $\mathcal{A}_2$ are two indep π-systems. Show that $\sigma(\mathcal{A}_1)$ and $\sigma(\mathcal{A}_2)$ also indep.
   (b) $\{S_n, \mathcal{F}_n\}$ is martingale. Show that $\{S_n, \sigma(S_n)\}$ is also a martingale.
   (c) iid X_i, if a.s. converge then X_i has finite mean.
   (d) W_n converge to normal in distribution. Show that $\sup_{x\in R}|P(W_n \leq x) - \Phi(x)|$ tends to 0.
2. (a) X_i are independent but not iid r.v., $0 \leq X_i \leq 1$. Show that $\frac{S_n}{E(S_n)} \to 1$ a.s.
   (b) Show that $B_n = Var(S_n)$ then $\frac{S_n}{B_n}$ converge to normal in distribution.
   (c) If $E(S_n)$ is finite when n tends to infinite, prove there exists a r.v. S that $S_n \to S$ a.s. with $ES_n \to ES$.
3. (a) X_i follow Cauchy, average of X_i follows what?
   (b) A_n and B_n converge in dist to A and a const $b > 0$. Show that A_nB_n converges to bA.
   (c) Wald's second identity.
4. (a) $P(X_k = k) = P(X_k = -k) = \frac{1}{2k^2}$, $P(X_k = 1) = P(X_k = -1) = (1 - ...)/2$. Show $\frac{S_n}{\sqrt{n}\log n} \to 0$ a.s.
   (b) $\lim Var(S_n)/n$.
   (c) limiting distribution of $S_n/\sqrt{Var(S_n)}$.

(d) limiting distribution of $S_n/\sqrt{\sum X_i^2}$.

5. (a) X_i follow unif(0,1), limiting distribution of $\sum 1\{X_i S_n \leq 1\}$.
   (b) Submartingale with finite optimum stopping time, $E(S_{\min(\tau,n)}) \leq E(S_n)$.
   (c) Doob's inequality.
   (d) $E(\max S_n^+)^2 \leq 4\sum E(X_i)^2$. X_i martingale diff.

4 Qualify 8/6/2018

4.1 # 5005

- $\xi_n \xrightarrow{d} \xi$ and $\eta_n \xrightarrow{d} \eta$, find a sufficient condition of $\xi_n + \eta_n \xrightarrow{d} \xi + \eta$
- X_i i.i.d. following U(-1,1).(1)find the limiting distribution of $M_n = \sum_{i<j\leq n} X_i X_j$. (2)prove $S_n/(\sqrt{n}logn) \to 0$ a.s.(3)prove $P(|S_n| > \sqrt{n}x) < 2e^{-\frac{x^2}{2}}$
- $\{S_n, F_n\}$ submartingale (1) N stopping time, prove $S_{min\{n,N\}} \leq S_n$ (2) $x > 0$, prove $P(max_{1\leq i\leq n} S_i^+ > x) \leq \frac{ES^+}{x}$ (3)prove $E(max_{1\leq i\leq n} S_i^2) \leq 4\sum_{i\leq n} EX_i^2$ $X_i = S_i - S_{i-1}$
  (We may have a typo here: we need $\{S_n, F_n\}$ is a martingale)

4.2 # 5010

- U(0,1) (1) $E(X_1/X_{(1)})^k$ (2) show $\sum(X_i - X_{(1)})$ and $X_{(1)}$ are independent.
- MLE: $(x_i, y_i) \sim Poisson(e^{\lambda_i}, e^{\lambda_i + W_i\beta})$. Impose conditions and show $\hat{\beta}$ has consistency and asymptotically normality.
- UMVUE for EX (1) $U(\theta - \tau, \theta + \tau)$ (2) unknowm distribution F (3) If (2) is contained in (1), explain why the UMVUE in (1) can't be UMVUE in (2).
- Test for $N(\theta_i, \sigma^2)$, $H_0 : \theta_i = 0$ versus $H_0 : \theta_i = \theta_{i0}$

4.3 # 5020

- 1.$[y|\mu, \Sigma] \sim N(\mu, \Sigma)$ (1)Assume Σ is known, give a conjugate prior of μ and prove the conjugacy. (2)Assume Σ is unknown, $\Sigma \sim IW(\Sigma_0, \phi_0)$ $[\mu|\Sigma] \sim N(\mu_0, \frac{\Sigma}{k_0})$. Derive the posterior joint distribution of μ, Σ

- 2 Mixture SEM. (1)List the assumptions of this model. (2)The challeges and stategies. (3)Detailed porcedures of estimation and model selection. (4)Consider examples in real life and explain what insights mixture model provide compared with convention regression.

4.4 # 5030

- 1. a linear model with a 3x2 design matix . parameter estimation and hypotesis testing(statistics, the critical region)

- 2.Ridge regression. Prove that there always exist $\lambda > 0$ under which the ridge estimator has smaller MSE compared with least square estimator.

No.

Date 2018/6/8 qualify exam: (类似期末)

probability: 1. ① $y_n \xrightarrow{d} \eta$, $\xi_n \xrightarrow{d} \xi$, $y_n+\xi_n \xrightarrow{d} \xi+\eta$? 加什么条件?

② $P(X_k = \pm k) = \frac{1}{2k^2}$, $P(X_k = \pm 1) = \frac{1}{2}(1-\frac{1}{k^2})$

$\begin{cases} S_n/(\sum_1^n X_i^2) \to ? & \text{(limiting distribution)} \\ S_n/\sqrt{n}\log n \to 0 \text{ a.s.} \end{cases}$

2. $X_i \overset{iid}{\sim} U(-1,1)$

(a). limiting distribution $W_n = \frac{1}{n}\sum_{1\le i<j\le n} X_i X_j$

(b) $S_n/\sqrt{n}\log n \to 0$ a.s.

(c). $P(|S_n| > \sqrt{n}t) \le 2\exp(-\frac{t^2}{2})$ [类似 self-normalized inequality]

3. martingale inequality. S_n submartingale.

①. $ES_{n\wedge T} \le ES_n$.

②. $xP(\max_{0\le m\le n} S_m^+ \ge x) \le E S_n^+ 1\{\max_{0\le m\le n} S_m^+ \ge x\}$.

③. $E(\max_{0\le m\le n} S_m^+)^2 \le 4\sum_1^n EX_i^2$

$S_n - S_{n-1} = X_n$..

SEM: Song

1. posterior $\begin{cases} \Sigma \text{ known} \\ \Sigma \text{ unknown} \quad IW. \end{cases}$

2. ①. Assumption of Mixture Model

②. What challenge will meet in Mixture model? How to Solve the problem?

③. How to estimate & model selection

④. Use real example to illustrate Mixture model & linear regression?

linear model.

1. $$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

a. $\hat{\theta}_1, \hat{\theta}_2$?

b. $\hat{\sigma}^2$? by (Y_1, Y_2, Y_3) & $(\hat{\theta}_1, \hat{\theta}_2)$

① Let $\ell(P)$ is the smallest λ-system containing
$\ell(P)$ is a σ-field. Thus $\sigma(P) \subset \ell(P) \subset \mathcal{L}$

$E(\max S_k^+ \wedge M)^p$ $p-1$ $\frac{1}{p}$

第＿＿＿＿＿題
（答題不得寫在紅線外）

第　頁

4^p

$P(S_n > 0$ ~~fail~~ for all $n \geq 1) > 0$.

No.

Date

c. Test $\theta_1 = 2\theta_2$, derive test statistic & rejection region with significance level α.

2. prove $\hat{\beta}$ridge $\exists \lambda$, s.t. $MSE(\hat{\beta}^{ridge}) < MSE(\hat{\beta}^{ls})$.

Inference 4選3

1. $U(0,\theta)$? $U(0,\theta)$.

① $E[X_1/X_{(n)}]^k$ compute.

② $\sum(X_i - X_{(1)}) \perp\!\!\!\perp X_{(1)}$

2. (X_i, Y_i) from poisson with mean $(e^{\lambda_i}, e^{\lambda_i + \beta W_i})$, W_i covariate.
prove MLE of β is consistant and asymptotic normal.

3. ① $X_i \overset{iid}{\sim} U(\theta - \tau, \theta + \tau)$. UMVUE for EX_1. 什 ∃

② $X_i \overset{iid}{\sim}$ unknown $F(\cdot)$ UMVUE for EX_1. 什 ∃

③ UMVUE in ① don't ~~suite~~ apply to ②

[different group]

4. $X_i \sim N(\mu_i, \sigma^2)$, σ^2 known, ump of level α

H_0: $\mu_i = \cdots = 0$ vs. H_1. $\mu_i = \theta_{i0}, \cdots$

find test statistics, rejection region.