QE 2022

Jun. 16

1 5005

1. $sup\ E|S_n| < +\infty$, $S_n \xrightarrow{a.s.} S$. Prove or give a counter example on $E|S_n - S| \to 0$
2. $\{X_n\}$ independent, $p_n = P(X_n = a_n) = 1 - P(X_n = -a_n)$, characterize the sequence $\{(a_n,\ p_n)\}$ such that $\sum X_n$ converges almost surely.
3. Simple random walk recurrent, prove that the probability to reach any given point b is 1.
4. Lindeberg randomized method. $E[X_i] = E[Y_i] = 0, \sum_{i=1}^{n} E[X_i^2] = \sum_{i=1}^{n} E[Y_i^2] = n, E[X_i^3], E[Y_i^3]$ bounded by γ, $W_n = \sqrt{n}\bar{X_n}, V_n = \sqrt{n}\bar{Y_n}$, prove that $E[g(W_n) - g(V_n)] \leq \frac{C\gamma||g'''||_\infty}{something}$. [Hint: write LHS as $(1/n!)\sum_{\sigma(n)}(g(W_{\sigma(n)} - g(V_{\sigma(n)})))$, where $\sigma(n)$ is the permutation of n variables.](Don't remember the exact right hand side. NOT sure about the hint.)

2 5010

1. $Poisson(\lambda)$, $\{X_n\}$ i.i.d.,
   - Give a complete and sufficient statistic.
   - Prove that $E[B^{n\bar{X_n}}] = e^{n(B-1)\lambda}$.
   - Give a UMVUE on $e^{-\lambda}$.
   - Is $e^{-\bar{X_n}}$ a consistent estimator of $e^{-\lambda}$?
   - Asymptotic normality of $e^{-\bar{X_n}}$.
2. $\theta \in [-C, C]$, $X \sim N(\theta, 1)$. $(C = \frac{1}{2})$
   - $Ctanh(CX)$ minimax (NOT sure).
   - MLE?
   - MLE's risk function?
   - Sketch two risk functions.
3. $f(x) = aexp(-a(x-b))$, $\{X_n\}$ i.i.d. Both parameters unknown. $a > 0$.
   - UMPT of $H_0 : b_0$ versus $H_1 : b \neq b_0$
   - UMPT of $H_0 : a_0, b_0$ versus $H_1 : a > a_0, b < b_0$. UMPT for $H_0 : a_0, b_0$ versus $H_1 : a \neq a_0, b \neq b_0$ exists?
4. (Bonus) Maximize $\prod_{i=1}^{n} p_i$ subject to $p_i > 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i\boldsymbol{u}_i = 0$. Prove that solution is: $\hat{p_i} = \frac{1}{n(1+\lambda^T u_i)}$, where λ satisfies $\sum_{i=1}^{n} \frac{u_i}{1+\lambda^T u_i} = 0$. Here, u_i and $\lambda \in R^s$ are vectors. (Not covered in the course, see Shao Jun exercise & solution book for *Mathematical Statistics* Exercise 5.7)

3 5020

1. Consider model

$$p(y_{ik}|\boldsymbol{\omega}_i) = exp\{[y_{ik}\vartheta - b(\vartheta_{ik})]/\psi_{\epsilon k} + c_k(y_{ik}, \phi_{\epsilon k})\},$$
$$E(y_{ik|\boldsymbol{\omega}_i}) = \dot{b}(\vartheta_{ik}), Var(y_{ik}|\boldsymbol{\omega}_i) = \psi_{\epsilon k}\ddot{b}(\vartheta_{ik}),$$

   y_{ik} is binary.

   - Specify $\vartheta_{ik}, b(\vartheta_{ik}), \psi_{\epsilon k}, C_k(y_{ik}, \psi_{\epsilon k}), \dot{b}(\vartheta_{ik})$.
   - Prior and derive posterior distribution.
   - Why should we use $p(\psi_{\epsilon k})$ and $p(\boldsymbol{\Lambda}_k|\psi_{\epsilon k})$ instead of $p(\psi_{\epsilon k})$ and $p(\boldsymbol{\Lambda}_k)$? Is there any conjugate prior for $p(\psi_{\epsilon k})$ and $p(\boldsymbol{\Lambda}_k|\psi_{\epsilon k})$?
   - When using MH algorithm, the proposal distribution for $p(\psi_{\epsilon k|}.)$ and $p(\boldsymbol{\Lambda}_k|\cdot)$ are $N(\cdot, \sigma_\psi^2 \Omega_{\psi k})$ and $N(\cdot, \sigma_\lambda^2 \boldsymbol{\Omega}_{\lambda k})$. Show that

$$\Omega_{\psi k}^{-1} = 1 - n/2 - \alpha_{0\epsilon k} - 2\sum_{i=1}^{n}[y_{ik}\vartheta_{ik} - b(\vartheta_{ik})] - \ddot{c_k}(y_{ik}, \psi_{\epsilon k}) + 2\beta_{0\epsilon k},$$

$$\boldsymbol{\Omega}_{\lambda k}^{-1} = \sum_{i=1}^{n} \ddot{b}(\vartheta_{ik})\boldsymbol{\omega}_i\boldsymbol{\omega}_i^T + \psi_{\epsilon k}^{-1} + \psi_{\epsilon k}^{-1}\boldsymbol{H}_{0yk}^{-1}.$$

   - How could we derive MLE($\boldsymbol{\theta}$) through CSA?
   - Explain why classical CSA approach cannot be applied to the analysis of advanced SEMs, such as nonlinear.

2. A path diagram for nonlinear SEM, all are continuous variables.
   - Write down the explicit form.
   - Assumption and identifiability.
   - How could we extend it to mixture, multi-level and multisample model? What is the difference in identifiability? How could we solve these?
   - Derive posterior distribution for each model.

4 5030

1. Design matrix

$$X = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & -1 \end{bmatrix}$$

   , $Y = X\boldsymbol{\theta} + \epsilon$, ϵ i.i.d. with mean 0 and variance σ^2

   - OLS.
   - Unbiased estimator for σ^2
   - Test $H_0 : \theta_1 = 2\theta_2$

2. Consider a regression model, $Y = X\beta + \epsilon$, where $X_{n\times p}$ is full column rank. $Y = (Y_1, ..., Y_n)'$. Further, assume that $Var(\epsilon) = \sigma^2 I$. Let e_i be the i-th residual, h_{ii} be the i-th diagonal element of the hat matrix. Let $\hat{\beta}$ and $\hat{\beta}_{(i)}$ be the LSE of β with and without the i-th case in the data respectively.

   - Prove $X'X = X'_{(i)}X_{(i)} + x_i x_i'$, where denote the regression matrix with the i-th row x_i deleted(should be a column vector).
   - Show that $(X'_{(i)}X_{(i)})^{-1} = (X'X)^{-1} + \frac{(X'X)^{-1}x_i x_i'(X'X)^{-1}}{1-h_{ii}}$
   - Prove that $\hat{\beta} - \hat{\beta}_{(i)} = \frac{(X'X)^{-1}x_i e_i}{1-h_{ii}}$.