Statistics 581, Problem Set 3 Solutions

Wellner; 10/15/2014

1. Ferguson, ACILST, page 34, problem 1(a) (modified slightly)
Suppose that $X_1, X_2, \ldots$ are i.i.d. in R^2 with distribution giving probability θ_1 to $(1,0)'$, probability θ_2 to $(0,1)'$, θ_3 to $(0,0)'$ and θ_4 to $(-1,-1)'$ where $\theta_j \ge 0$ for $j = 1,2,3,4$ and $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$.
(a) Find $\mu = E(X_1)$.
(b) Compute $E(X_1 X_1^T)$ and $\Sigma = E(X_1 - \mu)(X_1 - \mu)^T$.
(c) Find the limiting distribution of $\sqrt{n}(\overline{X}_n - \mu)$ and describe the resulting approximation to the distribution of $\overline{X}_n$.
(d) Find values of $(\theta_1, \ldots, \theta_4)$ such that Σ has rank 1 and $\det(\Sigma) = 0$.

Solution: (a) The mean of X_1 is given by

$$E(X_1) = \theta_1(1,0)' + \theta_2(0,1)' + \theta_3(0,0)' + \theta_4(-1,-1)' = (\theta_1 - \theta_4, \theta_2 - \theta_4)\,.$$

(b) Now

$$\begin{aligned} E(XX') &= \theta_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \theta_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \theta_3 \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \theta_4 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \theta_1 + \theta_4 & \theta_4 \\ \theta_4 & \theta_2 + \theta_4 \end{pmatrix}. \end{aligned}$$

Hence it follows that

$$\begin{aligned} \Sigma &= E(XX') - E(X)E(X') \\ &= \begin{pmatrix} \theta_1 + \theta_4 & \theta_4 \\ \theta_4 & \theta_2 + \theta_4 \end{pmatrix} - \begin{pmatrix} (\theta_1 - \theta_4)^2 & (\theta_1 - \theta_4)(\theta_2 - \theta_4) \\ (\theta_1 - \theta_4)(\theta_2 - \theta_4) & (\theta_2 - \theta_4)^2 \end{pmatrix} \\ &= \begin{pmatrix} \theta_1 + \theta_4 - (\theta_1 - \theta_4)^2 & \theta_4 - (\theta_1 - \theta_4)(\theta_2 - \theta_4) \\ \theta_4 - (\theta_1 - \theta_4)(\theta_2 - \theta_4) & \theta_2 + \theta_4 - (\theta_2 - \theta_4)^2 \end{pmatrix}. \end{aligned}$$

(c) By the multivariate CLT it follows that

$$\sqrt{n}(\overline{X}_n - E(X_1)) \to_d N_2(0, \Sigma)\,.$$

For example, if $\theta_j = 1/4$ for $j = 1,2,3,4$, then $E(X_1) = 0$,

$$\Sigma = \begin{pmatrix} 1/2 & 1/4 \\ 1/4 & 1/2 \end{pmatrix}$$

so the variances are both 1/2 and the correlation is $(1/4)/\sqrt{(1/2)(1/2)} = 1/2$; moreover $\det(\Sigma) = 1/4 - 1/16 = 3/16 > 0$. In this case the resulting normal approximation to the distribution of $\overline{X}_n$ is centered at 0 with a variance-covariance matrix $n^{-1}\Sigma$ with Σ as in the last display.

(d) Note that Σ does not depend explicitly on θ_3, but only through the constraint that $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$. When $\theta_3 = 1$, then $\theta_1 = \theta_2 = \theta_4 = 0$, $P(X_1 = 0) = 1$ and $\Sigma = 0$ has rank 0. Thus we reduce to the case $\theta_3 = 0$. If $\theta_2 = \theta_3 = 0$ and $\theta_1 = a = 1 - \theta_4$, then

$$\Sigma = \begin{pmatrix} 4a(1-a) & 2a(1-a) \\ 2a(1-a) & a(1-a) \end{pmatrix},$$

and $\det(\Sigma) = 0$. Thus for $\underline{\theta}) = (a, 0, 0, 1-a)$ with $0 < a < 1$, Σ has rank 1. Similarly Σ has rank 1 for $\theta = (0, a, 0, 1-a)$ with $0 < a < 1$.

2. Suppose that $X_1, X_2, \ldots$ are i.i.d. (μ, σ^2) with $\mu_4 < \infty$. Let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ be the sample mean and sample variance respectively.
(a) Show that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}.$$

(b) Suppose $\sigma > 0$. Use (a) to find the limiting distribution of the sample *signal-noise ratio* $D_n \equiv \overline{X}_n / S_n$; i.e. show that $\sqrt{n}(D_n - d) \to_d N(0, V^2)$ with $d \equiv \mu/\sigma$ and find V^2.

Solution: (a) Since $S_n^2 = n^{-1}\sum_{i=1}^n (X_i - \mu)^2 + o_p(1/\sqrt{n})$, we have

$$\begin{aligned} \sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \begin{pmatrix} X_i - \mu \\ (X_i - \mu)^2 - \sigma^2 \end{pmatrix} + o_p(1) \\ &\to_d \underline{Z} \sim N_2(0, \Sigma) \end{aligned}$$

by the multivariate CLT where Σ is as given above. Note that with $\underline{Y}_i \equiv (X_i - \mu, (X_i - \mu)^2 - \sigma^2)^T$, to apply the multivariate CLT we only need to verify that $E\|\underline{Y}_1\|^2 < \infty$ and compute

$$\Sigma_Y = E(\underline{Y}_1 \underline{Y}_1^T) = \begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}.$$

(b) The function $g(u, v) = u/\sqrt{v}$ is differentiable at points (u, v) with $v \neq 0$, and the derivative is $\nabla g(u,v) = (1/\sqrt{v}, u(-1/2)v^{-3/2})$ so that $\nabla g(\mu,\sigma^2) = (1/\sigma, (-1/2)\mu\sigma^{-3}) = (1/\sigma)(1, -(1/2)\mu/\sigma^2))$. Hence it follows from the delta method (g' theorem) that

$$\begin{aligned} \sqrt{n}(D_n - d) &= \sqrt{n}(g(\overline{X}_n, S_n^2) - g(\mu,\sigma^2)) \\ &\to_d \nabla g \cdot \underline{Z} \sim N(0, \nabla g^T \Sigma \nabla g) \end{aligned}$$

and it is easy to calculate that

$$\begin{aligned} \nabla g^T \Sigma \nabla g &= \frac{1}{\sigma^4}\left\{\sigma^4 - \mu\mu_3 + \frac{1}{4}d^2(\mu_4 - \sigma^4)\right\} \\ &= 1 - d\gamma_1 + \frac{1}{4}d^2(2+\gamma_2) \end{aligned}$$

where $\gamma_1 \equiv \mu_3/\sigma^3$ and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$. Note that when the X_i's are normal (so $\gamma_1 = \gamma_2 = 0$), this reduces to $1 + d^2/2$. Thus under normality we have

$$\sqrt{n}(g(D_n) - g(d)) \to_d N(0,1)$$

if $g(x) \equiv \sqrt{2}\text{arcsinh}(x/\sqrt{2})$.

3. Ferguson, ACILST, page 34, problem 1(b) (modified slightly)
Suppose that $X_1, \ldots, X_n$ is a sample from the Poisson distribution with parameter $\lambda > 0$: $P(X_1 = k) = \exp(-\lambda)\lambda^k/k!$, $k = 0, 1, \ldots$. Let $Z_n = (1/n)\sum_{i=1}^n 1_{[X_i=0]}$.
(a) What is the joint asymptotic distribution of

$$\sqrt{n}((\overline{X}_n, Z_n)' - (\lambda, e^{-\lambda})')?$$

(b) Let $p_0(\lambda) \equiv P_\lambda(X_1 = 0)$. What is the asymptotic distribution of $\hat{p}_0 \equiv p_0(\hat{\lambda}_n)$ where $\hat{\lambda}_n = \overline{X}_n$?
(c) What is the joint asymptotic distribution of $(Z_n, \hat{p}_0)$ (after centering and rescaling)?
(d) Compute the ratio of the asymptotic variances of the two estimators Z_n and $\hat{p}_0$ of $p_0(\lambda)$. Which estimator would you prefer if the Poisson model (assumption) holds? Which estimator would you prefer if the Poisson model (assumption) fails?

Solution: (a). Let $W_i \equiv (X_i, Y_i) \equiv (X_i, 1_{[X_i=1]})$. Then the W_i's are i.i.d. with mean $E(W_1) = (\lambda, e^{-\lambda})'$ and covariance matrix

$$\Sigma = \begin{pmatrix} \lambda & -\lambda e^{-\lambda} \\ -\lambda e^{-\lambda} & e^{-\lambda}(1 - e^{-\lambda}) \end{pmatrix}. \tag{1}$$

Hence the multivariate CLT implies that

$$\sqrt{n}(\overline{W} - E(W_1)) = \sqrt{n}((\overline{X}_n, Z_n)' - (\lambda, e^{-\lambda})) \to_d T \sim N_2(0, \Sigma) \tag{2}$$

where Σ is given in (1).

(b). Now $\hat{p}_0 = g(\overline{X}_n)$ where $g(v) = e^{-v}$. Hence $g'(v) = -e^{-v}$, $g'(\lambda) = -e^{-\lambda}$, and $\sqrt{n}(\overline{X}_n - \lambda) \to_d N(0, \lambda)$ by the CLT (or the first component of the convergence in distribution in part (a)). Hence it follows from the delta-method that

$$\sqrt{n}(\hat{p}_0 - p_0(\lambda)) = \sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d g'(\lambda)N(0, \lambda) = N(0, \lambda e^{-2\lambda}) .$$

(c). At this point it is a bit easier to study $(\hat{p}_0, Z_n) = g(\overline{X}_n, Z_n)$ where $g(u, v) \equiv (e^{-u}, v)$. Then in view of (2) and

$$\nabla g(\lambda, e^{-\lambda}) = \begin{pmatrix} -e^{-\lambda} & 0 \\ 0 & 1 \end{pmatrix} ,$$

it follows from the delta-method that

$$\sqrt{n}((\hat{p}_0, Z_n)' - e^{-\lambda}(1, 1)') \to_d \nabla g(\lambda, e^{-\lambda})T \sim N_2(0, \nabla g \Sigma (\nabla g)')$$

where

$$\nabla g \Sigma (\nabla g)' = \begin{pmatrix} \lambda e^{-2\lambda} & \lambda e^{-2\lambda} \\ \lambda e^{-2\lambda} & e^{-\lambda}(1 - e^{-\lambda}) \end{pmatrix} .$$

This is a situation in which we have two estimators of $P_\lambda(X_1 = 0) = p_0(\lambda)$, namely the MLE $\hat{p}_1 = p_0(\hat{\lambda})$ and the empirical (or "plug-in" estimator $Z_n = \#\{i \le n : X_i = 0\}/n$. Note that the ratio of the asymptotic variance of $\hat{p}_0$ to the asymptotic variance of Z_n is

$$ARE(\hat{p}_0, Z_n) \equiv \frac{\lambda e^{-2\lambda}}{e^{-\lambda}(1 - e^{-\lambda})} = \frac{\lambda e^{-\lambda}}{(1 - e^{-\lambda})} < 1$$

for all $\lambda > 0$. See the figure below

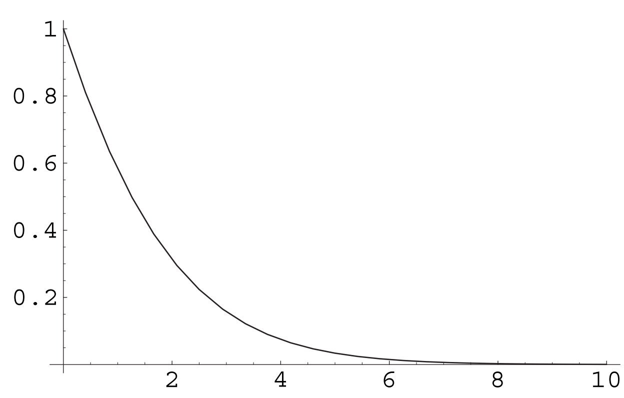

Figure 1: ARE of MLE relative to Plug-In.

4. A sequence of random variables Y_n is *bounded in probability* and we write $Y_n = O_p(1)$ if

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} P(|Y_n| > \lambda) = 0;$$

i.e. for each $\epsilon > 0$ there exist λ_ϵ and N_ϵ such that $P(|Y_n| > \lambda_\epsilon) < \epsilon$ for all $n > N_\epsilon$.
(a) Show that if $Y_n \to_d Y$ for some random variable Y, then Y_n is bounded in probability. (This is Lehmann and Casella, problem 8.24, page 77.)
(b) Give an example of a sequence of random variables Y_n that is bounded in probability, but does not converge in distribution.
(c) Lehmann and Casella, problem 8.25, page 77.
Solution: (a) Let F denote the limiting distribution, and let $Y \sim F$. Fix $\epsilon > 0$. Choose $\pm M \in C_F$ so large that $P(|Y| > M) \le 1 - F(M) + F(-M) < \epsilon/2$. Since $\pm M \in C_F$, and $Y_n \sim F_n$ converge in distribution to $Y \sim F$, we can find an $N = N_{M,\epsilon}$ so that $|F_n(\pm M) - F(\pm M)| < \epsilon/4$. Then for $n \ge N$ it follows that

$$\begin{aligned} P(|Y_n| > M) &= 1 - F_n(M) + F_n(-M-) \\ &\le 1 - F(M) + F(-M) + F(M) - F_n(M) + F_n(-M) - F(-M) \\ &< \frac{1}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon; \end{aligned}$$

i.e. Y_n is bounded in probability.
(b) Suppose that $Y_n = (-1)^n Y$ where $Y \sim$ Exponential(1). Then $|Y_n| \le |Y| = O_p(1)$, but $Y_{2n} \to Y$ while $Y_{2n+1} \to -Y$, and therefore $Y_n \not\to_d$.
(c) Problem 8.21 holds with o and O replaced by o_p and O_p:
(a') If $R_n = o_p(1/k_n)$, then $R_n/(1/k_n) = k_n R_n = o_p(1)$, so $k_n R_n = O_p(1)$, and hence $R_n = O_p(1/k_n)$.
(b') If $R_n = O_p(1)$ then R_n is bounded in probability. This is just the definition of R_n being $O_p(1)$.
(c') If $R_n = o_p(1)$, then $R_n \to_p 0$; this is just the definition of $R_n = o_p(1)$.
(d') If $R_n = O_p(1/k_n)$, and $k_n'/k_n \to \rho \in (-\infty, \infty)$, then $R_n = O_p(1/k_n')$: for every $\epsilon > 0$ there exists $M = M_\epsilon$ and $N = N_\epsilon$ such the for $n \ge N_\epsilon$ we have

$$P(|k_n R_n| > M) < \epsilon$$

and, perhaps by choosing N_ϵ somewhat larger, $|k_n'/k_n - \rho| < \epsilon$. Thus with $M' = M(|\rho| + \epsilon)$,

$$P(|k_n' R_n| > M') = P(|(k_n'/k_n) k_n R_n| > M') \le P((|\rho|+\epsilon)|k_n R_n| > M') \le P\left(|k_n R_n| > M\right) < \epsilon.$$

Problem 8.22 holds with o and O replaced by o_p and O_p:
(a') If $R_n = O_p(1/k_n)$ and $R_n' = O_p(1/k_n)$, then $R_n + R_n' = O_p(1/k_n)$: by the hypothesis, for each $\epsilon > 0$ there exist $M = M_\epsilon, M_\epsilon'$ such that for $n \ge N_\epsilon$

$$P(|k_n R_n| > M) < \epsilon \qquad \text{and} \qquad P(|k_n R_n'| > M') < \epsilon.$$

Let $\tilde{M} \equiv M_{\epsilon/2} + M'_{\epsilon/2}$. Then

$$P(|k_n(R_n + R'_n)| > \tilde{M}) \le P(|k_n R_n| > M_{\epsilon/2}) + P(|k_n R'_n| > M'_{\epsilon/2}) < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus $R_n + R'_n = O_p(1/k_n)$.
(b') If $R_n = o_p(1/k_n)$ and $R'_n = o_p(1/k_n)$, then $k_n R_n \to_p 0$ and $k_n R'_n \to_p 0$, so $k_n(R_n + R'_n) \to_p 0$; i.e. $R_n + R'_n = o_p(1/k_n)$.
Problem 8.23 holds with o and O replaced by o_p and O_p: Suppose that $k'_n/k_n \to \infty$.
(a') If $R_n = O_p(1/k_n)$ and $R'_n = O_p(1/k'_n)$, then $R_n + R'_n = O_p(1/k_n)$: The hypotheses imply that $k_n R_n = O_p(1)$ and $R'_n = O_p(1/k'_n)$. Thus

$$k_n(R_n + R'_n) = k_n R_n + \frac{k_n}{k'_n} k'_n R'_n = O_p(1) + o_p(1)O_p(1) = O_p(1);$$

i.e. $R_n + R'_n = O_p(1/k_n)$.
(b') If $R_n = o_p(1/k_n)$ and $R'_n = o_p(1/k'_n)$, then $R_n + R'_n = o_p(1/k_n)$: This follows easily since

$$k_n(R_n + R'_n) = k_n R_n + \frac{k_n}{k'_n} k'_n R'_n \to_p 0 + 0 \cdot 0 = 0.$$

5. Suppose that X is a random variable with finite fourth moment; $E|X|^4 < \infty$. Then $\mu_4 = E(X-\mu)^4$ is the fourth central moment of X. The ratio $\mu_4/\sigma^4 \equiv \kappa$ is the *kurtosis* of X (or of the distribution function F of X), and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*; note that for any $N(\mu, \sigma^2)$ random variable, $\gamma_2 = 0$. Investigate the value of γ_2 for various classical distributions (t_r, uniform, bernoulli, Poission(λ), ...). How big can γ_2 be? How small can γ_2 be?
Solution: Note that $\mu_4^{1/4} = \{E(X-\mu)^4\}^{1/4} \ge \{E(X-\mu)^2\}^{1/2} = \sigma$ by Liapunov's inequality. Thus $\mu_4/\sigma^4 \ge 1$ always, or $\gamma_2 \equiv \mu_4/\sigma^4 \ge -2$ with equality if $X = \pm 1$ with probability 1/2 each: then $\mu = 0$, $\sigma^2 = 1$, $\mu_4 = 1$, and $\gamma_2 = -2$.
For $X \sim N(0,1)$, $\gamma_2 = 0$ since $EX^4 = 3$.
For $X \sim t_r$, $r > 4$, $\gamma_2 = 6/(r-4) \nearrow \infty$ as $r \searrow 4$; $\gamma_2 \searrow 0$ as $r \nearrow \infty$.
For $X \sim$ Gamma(α, β), $\gamma_2 = 6/\alpha \nearrow \infty$ as $\alpha \searrow 0$.
For $X \sim$ Poisson(λ), $\gamma_2 = 1/\lambda \nearrow \infty$ as $\lambda \searrow 0$.
For $X \sim$ Bernoulli(p), $\gamma_2 = (1-p)^2/p + p^2/(1-p) - 3$ which $= -2$ when $p = 1/2$, and $\nearrow \infty$ when $p \to 0, 1$.

6. **Optional Bonus Problem 1.** Suppose that $X_1, X_2, \ldots$ are i.i.d. positive random variables, and define $\overline{X}_n \equiv n^{-1}\sum_{i=1}^n X_i$, $H_n \equiv 1/(n^{-1}\sum_{i=1}^n (1/X_i))$, and $G_n \equiv \{\prod_{i=1}^n X_i\}^{1/n}$ to be the *arithmetic, harmonic, and geometric* means respectively. We know that $\overline{X}_n \to_{a.s.} E(X_1) = \mu$ if and only if $E|X_1| < \infty$.
(a) Use the SLLN together with appropriate additional hypotheses to show that $H_n \to_{a.s.} 1/\{E(1/X_1)\} \equiv h$, and $G_n \to_{a.s.} \exp(E\{\log X_1\}) \equiv g$.
(c) Use the multivariate CLT and the delta method to find the joint limiting

distribution of $\sqrt{n}(\overline{X}_n - \mu, H_n - h, G_n - g)$. You will need to impose or assume additional moment conditions to be able to prove this. Specify these additional assumptions carefully.

Solution: (a) If $0 < E(1/X_1) < \infty$, then

$$\frac{1}{n}\sum_{i=1}^{n}(1/X_i) \rightarrow_{a.s.} E(1/X_1) > 0.$$

If $E|\log(X_1)| < \infty$, then

$$\log G_n = \frac{1}{n}\sum_{i=1}^{n}\log(X_i) \rightarrow_{a.s.} E\log X_1.$$

Thus by the continuous mapping theorem if both $E(1/X_1) < \infty$ and $E|\log X_1| < \infty$, it follows that

$$(H_n, G_n) \rightarrow_{a.s.} (1/E(1/X_1), \exp(E\log X_1)) \equiv (h, g).$$

(c) By the multivariate CLT, if $EX_1^2 < \infty$, $E(1/X_1)^2 < \infty$, and $E(\log X_1)^2 < \infty$, then

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ \overline{X^{-1}}_n - E(1/X_1) \\ \overline{\log X}_n - E\log X_1 \end{pmatrix} \rightarrow_d \underline{Z} \sim N_3(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} Var(X_1) & Cov(X_1, 1/X_1) & Cov(X_1, \log(X_1)) \\ Cov(X_1, 1/X_1) & Var(1/X_1) & Cov(1/X_1, \log X_1) \\ Cov(X_1, \log(X_1)) & Cov(1/X_1, \log X_1) & Var(\log(X_1)) \end{pmatrix}.$$

Hence by the delta method with $g(x, y, z) = (x, 1/y, \exp(z))$ so that $\nabla g(x, y, z) = \text{diag}(1, -y^{-2}, \exp(z))$ and $\nabla g(\mu, E(1/X_1), E(\log X_1)) = \text{diag}(1, -h^2, g)$, it follows that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ H_n - h \\ G_n - g \end{pmatrix} \rightarrow_d \nabla g \cdot \underline{Z} \sim N_3(0, \nabla g \Sigma \nabla g^T).$$

7. **Optional bonus problem 2:** (i) Ferguson, ACILST, page 34, problem 3.
(ii) Ferguson, ACILST, page 34, problem 4.

Solution: (i) If $X_1, X_2, \ldots$ are i.i.d. with $E(X_1^2) < \infty$, let $X_{ni} \equiv X_i - \mu$ for $i = 1, \ldots, n$. Then we have $E(X_{ni}) = 0$ and $\sigma_{ni}^2 = Var(X_1) = \sigma^2$, so that $\sigma_n^2 \equiv \sum_{i=1}^n \sigma_{ni}^2 = n\sigma^2$. The Lindeberg condition becomes, since the $Y_i \equiv X_i - \mu$ are i.i.d.,

$$\begin{aligned}
\frac{1}{\sigma_n^2} \sum_{i=1}^n E\{|X_{ni}|^2 1_{[|X_{ni}| \geq \epsilon \sigma_n]}\} &= \frac{1}{n\sigma^2} \sum_{i=1}^n E\{Y_1^2 1_{[|Y_1| \geq \epsilon \sigma \sqrt{n}]}\} \\
&= \frac{1}{\sigma^2} E\{Y_1^2 1_{[|Y_1| \geq \epsilon \sigma \sqrt{n}]}\} \to 0
\end{aligned}$$

by the dominated convergence theorem since $E(Y_1^2) = E(X_1 - \mu)^2 < \infty$. Thus the Lindeberg-Feller central limit theorem implies that

$$\frac{S_n}{\sigma_n} = \frac{\sum_{i=1}^n (X_i - \mu)}{\sigma \sqrt{n}} \to_d Z \sim N(0, 1).$$

But this can be rewritten as $\sqrt{n}(\overline{X}_n - \mu) \to_d \sigma Z \sim N(0, \sigma^2)$, so the usual (Lindeberg) CLT for i.i.d. real-valued random variables holds.

(ii) Let $P(X_j = \pm v_j) = p_j/2$ and $P(X_j = 0) = 1 - p_j$ for $j = 1, 2, \ldots$. Then $E(X_j) = 0$ and $Var(X_j) = E(X_j^2) = v_j^2 p_j = 1$ if $p_j = 1/v_j^2$. Let $v_j = j^{1/2}$, so that $p_j = 1/j$ for $j = 1, 2, \ldots$. Then $\sigma_{nj}^2 = 1$ for all j and $\sigma_n^2 = \sum_{j=1}^n \sigma_{nj}^2 = n$. Thus the Lindeberg condition becomes

$$\begin{aligned}
L_n(\epsilon) &= n^{-1} \sum_{j=1}^n E(X_j^2 1_{[|X_j| \geq \epsilon \sqrt{n}]} = n^{-1} \sum_{j=1}^n v_j^2 p_j 1_{[v_j \geq \epsilon \sqrt{n}]} \\
&= n^{-1} \sum_{j=1}^n 1_{[j \geq \epsilon^2 n]} \approx n^{-1}(n - \epsilon^2 n) \\
&\to 1 - \epsilon^2 \neq 0.
\end{aligned}$$

Furthermore $\sigma_{nj}^2 / \sigma_n^2 = n^{-1} \to 0$, so the Lindeberg condition becomes equivalent to asymptotic normality of $\sqrt{n}\overline{X}_n$. Since the Lindeberg condition fails, asymptotic normality of $\sqrt{n}\overline{X}_n$ necessarily fails as well.