Indian Statistical Institute, Kolkata

Large Sample Statistical Methods

Subhrajyoty Roy
Roll: MB1911

June 29, 2020

Contents

1 Problem 1

Problem. Let, F be a distribution function on $\mathbb{R}$ and $0 < p < 1$. Show that a real number c is the unique quantile of order p of F if and only if,

$$F(c - \epsilon) < p < F(c + \epsilon) \qquad \forall \epsilon > 0$$

Solution:

If part:

Here we assume, $F(c-\epsilon) < p < F(c+\epsilon)$ for any $\epsilon > 0$. Then we need to show that c is the unique quantile of order p of F.

Taking a sequence of $\epsilon_n \to 0$, and using the above inequality for ϵ_n instead of ϵ, we have;

$$F(c-) = \lim_{n\to\infty} F(c - \epsilon_n) \leq p \leq \lim_{n\to\infty} F(c + \epsilon_n) = F(c+)$$

However, since F is a distribution function, hence it is right continuous, $F(c+) = F(c)$. This reduces to,

$$F(c-) \leq p \leq F(c)$$

showing that c is a quantile of F of order p. Now, to show that c is the unique quantile, assume c' is another quantile of F of order p. Then we have by definition of quantile,

$$F(c'-) \leq p \leq F(c') \tag{1}$$

If $c > c'$, then $\epsilon = (c - c') > 0$, and by the given condition $F(c') = F(c - \epsilon) < p$ which contradicts (1). On the other hand, if $c' > c$, then $\epsilon = (c' - c) > 0$. Then again by the given condition, $F(c' - \epsilon/2) = F(c + \epsilon/2) > p$, but since $F(c'-) \leq p$, it must be the case that $F(c' - \epsilon/2) \leq F(c'-) \leq p$, leading to another contradiction.

Only if part: Now, we assume that c is the unique quantile of order p of F, and we need to show that for any given $\epsilon > 0$, we have; $F(c - \epsilon) < p < F(c + \epsilon)$.

Now since c is a p-th quantile, by definition we have $F(c-) \leq p \leq F(c)$. Also, since F is non-decreasing, we have for any given $\epsilon > 0$, $F(c - \epsilon) \leq F(c-) \leq p \leq F(c) \leq F(c + \epsilon)$. We simply need to show that it holds with a strict inequality.

Let us assume otherwise. Then we consider two cases.

1. First case is that for some $\epsilon_0 > 0$, $F(c - \epsilon_0) = p$. Then clearly, $F((c - \epsilon_0)-) \leq p \leq F(c - \epsilon_0)$, therefore, showing that $(c - \epsilon_0)$ is also a p-th order quantile. But, as $\epsilon_0 > 0$, $c \neq (c - \epsilon_0)$, contradicting the assumption that c is the unique p-th quantile.

2. For the second case, assume for some $\epsilon_0 > 0$, $F(c + \epsilon_0) = p$. Then again by exactly same logic, $(c + \epsilon_0)$ is another p-th quantile, contradicting to the uniqueness of c.

Therefore, we must have, $F(c - \epsilon) < p < F(c + \epsilon)$ for any $\epsilon > 0$, completing the proof.

□

2 Problem 2

Problem. Let, $X_1, X_2, \dots X_n$ be a random sample from a distribution symmetric about 0, and having finite 8-th moment. Find the asymptotic distribution of the following measures of skewness and kurtosis:

(a) $g_{1n} = \dfrac{m_{3n}}{m_{2n}^{3/2}}$

(b) $g_{2n} = \dfrac{m_{4n}}{m_{2n}^{2}} - 3$

where m_{rn} is the r-th order sample central moment. First find the result for the general case and then use it to derive for the case when the underlying distribution is $N(\mu, \sigma^2)$.

Solution: By a simple application of multivariate central limit theorem for $X_1, X_2, \dots X_n$, and using **Theorem D** (in the section related to asymptotic distribution of moments), we have $\sqrt{n}(T_n - \boldsymbol{\mu})$ is asymptotically normal with mean parameter $\mathbf{0}$ and dispersion matrix $\Sigma = ((\sigma_{ij}))$, where;

$$T_n = \begin{bmatrix} m_{1n} \\ m_{2n} \\ m_{3n} \\ m_{4n} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix}$$

and

$$\sigma_{ij} = \mu_{i+j} - \mu_i\mu_j - i\mu_{i-1}\mu_{i+1} - j\mu_{j-1}\mu_{j+1} + ij\mu_{i-1}\mu_{j-1}\mu_2$$

where m_{rn} is the r-th order sample central moment and μ_r is the r-th order population central moment.

Now, since $X_1, X_2, \dots X_n$ is a random sample from a distribution symmetric about 0, clearly, all odd order population central moments will be equal to 0, i.e. $\mu_1 = \mu_3 = 0$. Therefore, $\boldsymbol{\mu} = (0, \mu_2, 0, \mu_4)^T$ Also, let us denote the set of permissible values for the vector T_n as A, i.e.

$$A = \{(x_1, x_2, x_3, x_4) : x_i \in \mathbb{R} \quad \forall i = 1, 2, 3, 4 \text{ and } x_2 \geq 0, x_4 \geq 0\}$$

(a) Consider the function, $g : A \to \mathbb{R}$ given by;

$$g(x_1, x_2, x_3, x_4) = \frac{x_3}{x_2^{3/2}}$$

$$\therefore \quad \nabla g(\boldsymbol{x}) = \left(0, -\frac{3}{2}\frac{x_3}{x_2^{5/2}}, \frac{1}{x_2^{3/2}}, 0\right)$$
$$\Rightarrow \quad \nabla g(\boldsymbol{\mu}) = \left(0, 0, \mu_2^{-3/2}, 0\right)$$

Now, applying delta method, we get that, $\sqrt{n}(g(T_n) - g(\boldsymbol{\mu}))$ is asymptotically normal with mean $\mathbf{0}$ and dispersion $\nabla g(\boldsymbol{\mu})^T \Sigma \nabla g(\boldsymbol{\mu})$. But note that, $g(T_n) = g_{1n}$ and $g(\boldsymbol{\mu}) = \dfrac{\mu_3}{\mu_2^{3/2}} = 0$. And finally,

$$\begin{aligned}
\nabla g(\boldsymbol{\mu})^T \Sigma \nabla g(\boldsymbol{\mu}) &= \frac{1}{\mu_2^3}\sigma_{33} \\
&= \frac{1}{\mu_2^3}\left(\mu_6 - \mu_3^2 - 6\mu_2\mu_4 + 9\mu_2^3\right) \\
&= \frac{1}{\mu_2^3}\left(\mu_6 - 6\mu_2\mu_4 + 9\mu_2^3\right) \\
&= \frac{\mu_6}{\mu_2^3} - 6\frac{\mu_4}{\mu_2^2} + 9
\end{aligned}$$

So, we have;

$$\sqrt{n}g_{1n} \text{ is Asymptotically Normal} \left(0, \frac{\mu_6}{\mu_2^3} - 6\frac{\mu_4}{\mu_2^2} + 9\right)$$

In case the samples $X_1, X_2, \dots X_n$ comes from a normal distribution with parameters μ and σ^2, we consider the standardized samples, $Z_i = \dfrac{(X_i - \mu)}{\sigma}$. Then, $Z_1, Z_2, \dots Z_n \sim N(0,1)$ and more importantly, the sample and population skewness and kurtosis coefficients of X_i's and Z_i's are same. Hence, it is enough to consider Z_i's instead X_i's, i.e. the underlying distribution being the standard normal distribution and the same asymptotic distribution can be applied in any $N(\mu, \sigma^2)$ distribution for any $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

Considering standard normal distribution, we have;

$$\mu_{2k} = (2k-1)(2k-3)\dots 3.1 \qquad \forall k \in \mathbb{N}$$

Using this, we get $\mu_6 = 15, \mu_4 = 3, \mu_2 = 1$, and hence, $\dfrac{\mu_6}{\mu_2^3} - 6\dfrac{\mu_4}{\mu_2^2} + 9 = (15 - 18 + 9) = 6$. Therefore, in case the underlying distribution is a normal distribution,

$$\sqrt{n}g_{1n} \text{ is Asymptotically Normal}\,(0, 6)$$

(b) For this case, consider the function, $h : A \to \mathbb{R}$ given by;

$$h(x_1, x_2, x_3, x_4) = \frac{x_4}{x_2^2} - 3$$

$$\begin{aligned}
\therefore \quad \nabla h(\boldsymbol{x}) &= \left(0, -2\frac{x_4}{x_2^3}, 0, \frac{1}{x_2^2}\right) \\
\Rightarrow \quad \nabla h(\boldsymbol{\mu}) &= \left(0, -2\frac{\mu_4}{\mu_2^3}, 0, \frac{1}{\mu_2^2}\right)
\end{aligned}$$

Now, applying delta method, we get that, $\sqrt{n}(h(T_n) - h(\boldsymbol{\mu}))$ is asymptotically normal with mean $\mathbf{0}$ and dispersion $\nabla h(\boldsymbol{\mu})^T \Sigma \nabla h(\boldsymbol{\mu})$. But note that, $h(T_n) = g_{2n}$ and $h(\boldsymbol{\mu}) = \dfrac{\mu_4}{\mu_2^2} - 3$, the population coefficient of kurtosis. Hence,

$$
\begin{aligned}
\nabla h(\boldsymbol{\mu})^T \Sigma \nabla h(\boldsymbol{\mu}) &= 4\frac{\mu_4^2}{\mu_2^6}\sigma_{22} - 4\frac{\mu_4}{\mu_2^5}\sigma_{24} + \frac{1}{\mu_2^2}\sigma_{44} \\
&= 4\frac{\mu_4^2}{\mu_2^6}\left[\mu_4 - \mu_2^2 - 4\mu_1\mu_3 + 4\mu_1^2\mu_2\right] - 4\frac{\mu_4}{\mu_2^5}\left[\mu_6 - \mu_2\mu_4 - 2\mu_1\mu_3 - 4\mu_3\mu_5 + 8\mu_1\mu_3\mu_2\right] \\
&\qquad + \frac{1}{\mu_2^2}\left[\mu_8 - \mu_4^2 - 8\mu_3\mu_5 + 16\mu_3^2\mu_2\right] \\
&= 4\frac{\mu_4^2}{\mu_2^6}\left(\mu_4 - \mu_2^2\right) - 4\frac{\mu_4}{\mu_2^5}\left(\mu_6 - \mu_2\mu_4\right) + \frac{1}{\mu_2^2}\left(\mu_8 - \mu_4^2\right)
\end{aligned}
$$

Therefore, again by application of delta method, we simply have,

$$
\sqrt{n}(g_{2n} - \kappa) \text{ is Asymptotically Normal } \left(0, 4\frac{\mu_4^2}{\mu_2^6}\left(\mu_4 - \mu_2^2\right) - 4\frac{\mu_4}{\mu_2^5}\left(\mu_6 - \mu_2\mu_4\right) + \frac{1}{\mu_2^2}\left(\mu_8 - \mu_4^2\right)\right)
$$

where $\kappa = \dfrac{\mu_4}{\mu_2^2} - 3$, the population coefficient of kurtosis.

In case the underlying distribution of the samples is standard normal distribution, using $\mu_8 = (7 \times 5 \times 3) = 105, \mu_6 = 15, \mu_4 = 3$ and $\mu_2 = 1$, and noting that $\kappa = 0$, we have;

$$
\begin{aligned}
& 4\frac{\mu_4^2}{\mu_2^6}\left(\mu_4 - \mu_2^2\right) - 4\frac{\mu_4}{\mu_2^5}\left(\mu_6 - \mu_2\mu_4\right) + \frac{1}{\mu_2^2}\left(\mu_8 - \mu_4^2\right) \\
= \;& 4 \times 3^2 \times (3 - 1^2) - 4 \times 3 \times (15 - 3) + (105 - 3^2) \\
= \;& 72 - 144 + 105 - 9 = 24
\end{aligned}
$$

i.e. for the case where underlying distribution is a normal distribution (or standard normal distribution, since similar to previous logic we can consider Z_i's here),

$$
\sqrt{n} g_{2n} \text{ is Asymptotically Normal } (0, 24)
$$

$\square$

3 Problem 3

Problem. Let, $X_1, X_2, \dots X_n$ be i.i.d $N(\theta, 1)$ variables where it is known that $|\theta| \le 1$. Find the mle of θ and also find its asymptotic distribution under any $\theta_0 \in (-1, 1)$.

Solution: We are given, $X_i \sim N(\theta, 1)$ be iid random variables, $i = 1, 2, \dots n$. Therefore, the likelihood function upto a constant multiple is given as;

$$\mathcal{L}(\theta) \propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(X_i - \theta)^2\right\}\mathbf{1}_{|\theta|\le 1} \tag{2}$$

Now to find the maximizer of the likelihood function $\mathcal{L}(\theta)$, letting $\overline{X}_n$ as the sample mean based on $X_1, X_2, \dots X_n$, consider the following;

$$\begin{aligned}
&\mathcal{L}(\theta) \propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(X_i - \overline{X}_n + \overline{X}_n - \theta)^2\right\}\mathbf{1}_{|\theta|\le 1} \\
\Rightarrow\ &\mathcal{L}(\theta) \propto \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + n(\overline{X}_n - \theta)^2\right]\right\}\mathbf{1}_{|\theta|\le 1} \\
\Rightarrow\ &\mathcal{L}(\theta) \propto \exp\left\{-\frac{n}{2}(\overline{X}_n - \theta)^2\right\}\mathbf{1}_{|\theta|\le 1}
\end{aligned}$$

Clearly, this likelihood as a function of θ decreases to both side of $\overline{X}_n$ as $|\overline{X}_n - \theta|$ increases in magnitude, and if θ lies outside the interval $[-1, 1]$, the likelihood function becomes 0.

Hence, the likelihood function is maximized at $\theta = \overline{X}_n$, if $\overline{X}_n \in [-1, 1]$, and otherwise, the maximization happens at the endpoint of the interval $[-1, 1]$ which is closest to $\overline{X}_n$. Therefore, the mle of θ is;

$$\widehat{\theta}_n = \begin{cases} \overline{X}_n & \text{if } (-1) \le \overline{X}_n \le 1 \\ 1 & \text{if } \overline{X}_n > 1 \\ (-1) & \text{if } \overline{X}_n < (-1) \end{cases}$$

Now, to find the asymptotic distribution of $\widehat{\theta}_n$, we consider the following quantity;

$$\begin{aligned}
&P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \le x\right) \\
= \ &P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \le x, \overline{X}_n \in (-1, 1)\right) + P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \le x, \overline{X}_n \ge 1\right) \\
&\quad + P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \le x, \overline{X}_n \le (-1)\right) \\
= \ &P_{\theta_0}\left(\sqrt{n}(\overline{X}_n - \theta_0) \le x\right)P_{\theta_0}\left(\overline{X}_n \in (-1, 1)\right) + P_{\theta_0}\left(\sqrt{n}(\delta_1 - \theta_0) \le x\right)P_{\theta_0}\left(\overline{X}_n \ge 1\right) \\
&\quad + P_{\theta_0}\left(\sqrt{n}(\delta_{-1} - \theta_0) \le x\right)P_{\theta_0}\left(\overline{X}_n \le (-1)\right)
\end{aligned}$$

where δ_x is the degenerate distribution at x. Now note that, $\overline{X}_n \sim N(\theta_0, \frac{1}{n})$ under θ_0 being the true parameter. Therefore,

$$\begin{aligned} P_{\theta_0}\left(\overline{X}_n < 1\right) &= \Phi\left(\sqrt{n}(1-\theta_0)\right) \\ P_{\theta_0}\left(\overline{X}_n < (-1)\right) &= \Phi\left(\sqrt{n}(-1-\theta_0)\right) \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution. Also note that, since the true parameter $\theta_0 \in (-1, 1)$, $(1-\theta_0) > 0$ and $(-1-\theta_0) < 0$. Hence, as $n \to \infty$,

$$\begin{aligned} \sqrt{n}(1-\theta_0) &\to \infty \\ \sqrt{n}(-1-\theta_0) &\to -\infty \end{aligned}$$

Hence, for sufficiently large n, i.e. for any $n \geq N = \max\left\{\frac{x^2}{(1-\theta_0)^2}, \frac{x^2}{(-1-\theta_0)^2}\right\}$, we have;

$$\begin{aligned} P_{\theta_0}(\sqrt{n}(\delta_1 - \theta_0) \leq x) &= 0 \\ P_{\theta_0}(\sqrt{n}(\delta_{-1} - \theta_0) \leq x) &= 1 \end{aligned}$$

Hence for any $n \geq N$, we have;

$$P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \leq x\right) = \Phi(x)\left[\Phi\left(\sqrt{n}(1-\theta_0)\right) - \Phi\left(\sqrt{n}(-1-\theta_0)\right)\right] + \Phi\left(\sqrt{n}(-1-\theta_0)\right)$$

Now taking limit as $n \to \infty$, we obtain;

$$\begin{aligned} & P_{\theta_0}\left(\sqrt{n}(\widehat{\theta}_n - \theta_0) \leq x\right) \\ = \;& \Phi(x)\left[\Phi\left(\sqrt{n}(1-\theta_0)\right) - \Phi\left(\sqrt{n}(-1-\theta_0)\right)\right] + \Phi\left(\sqrt{n}(-1-\theta_0)\right) \\ \xrightarrow{n\to\infty} \;& \Phi(x)\left[\Phi(\infty) - \Phi(-\infty)\right] + \Phi(-\infty) \\ = \;& \Phi(x) \end{aligned}$$

Therefore, for all $x \in \mathbb{R} = \mathcal{C}(\Phi)$, i.e. the continuity points of $\Phi(\cdot)$, the distribution function of $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ converges to the distribution function of standard normal distribution, i.e. we have,

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{d} Z \sim N(0, 1)$$

i.e. the suitably centered and scaled random variable of the MLE, $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ follows an asymptotic normal distribution with mean 0 and variance 1. $\square$

4 Problem 4

Problem. Let, $X_1, X_2, \ldots X_n$ be a random sample from a distribution with a density $f(x \mid \theta), \theta \in \Theta$, an open interval in $\mathbb{R}$. Assume suitable regularity conditions on the densities so that there is a consistent solution $\widehat{\theta}_n$ of the likelihood equation for which $\sqrt{n}(\widehat{\theta}_n - \theta)$ is $AN(0, \mathcal{I}^{-1}(\theta))$, under θ. Fix, $\theta_0 \in \Theta$ and set,

$$T_n = \begin{cases} \widehat{\theta}_n & \text{if } |\widehat{\theta}_n - \theta_0| > n^{-1/4} \\ \theta_0 & \text{if } |\widehat{\theta}_n - \theta_0| \leq n^{-1/4} \end{cases}$$

Find the asymptotic distribution of $\sqrt{n}(T_n - \theta)$ under θ.

Solution: We shall consider two separate cases, where true $\theta = \theta_0$ and another case where $\theta \neq \theta_0$.

Case 1: $\theta \neq \theta_0$.

Consider the quantity, $Z_n = \sqrt{n}(T_n - \widehat{\theta}_n)$. Note that,

$$Z_n = \begin{cases} 0 & \text{if } |\widehat{\theta}_n - \theta_0| > n^{-1/4} \\ \sqrt{n}(\theta_0 - \widehat{\theta}_n) & \text{if } |\widehat{\theta}_n - \theta_0| \leq n^{-1/4} \end{cases}$$

Therefore,

$$\begin{aligned} P_\theta(Z_n \neq 0) &\leq P_\theta\left(|\widehat{\theta}_n - \theta_0| \leq n^{-1/4}\right) \\ &= P_\theta\left(\widehat{\theta}_n \geq \theta_0 - n^{-1/4}, \widehat{\theta}_n \leq \theta_0 + n^{-1/4}\right) \\ &= P_\theta\left((\widehat{\theta}_n - \theta) \geq (\theta_0 - \theta) - n^{-1/4}, (\widehat{\theta}_n - \theta) \leq (\theta_0 - \theta) + n^{-1/4}\right) \end{aligned}$$

Now, if $\theta_0 > \theta$, then for sufficiently large n, we have, $(\theta_0 - \theta) - n^{-1/4} > \epsilon > 0$ for some $\epsilon > 0$. Hence in this case,

$$\begin{aligned} P_\theta(Z_n \neq 0) &\leq P_\theta\left((\widehat{\theta}_n - \theta) \geq (\theta_0 - \theta) - n^{-1/4}\right) \\ &\leq P_\theta\left((\widehat{\theta}_n - \theta) \geq \epsilon\right) \\ &\qquad \text{for sufficently large } n \text{ as,} \\ &\qquad (\widehat{\theta}_n - \theta) \geq ((\theta_0 - \theta) - n^{-1/4}) \implies (\widehat{\theta}_n - \theta) \geq \epsilon \\ &\leq P_\theta\left(|\widehat{\theta}_n - \theta| \geq \epsilon\right) \\ &\qquad \text{since, } (\widehat{\theta}_n - \theta) \geq \epsilon \implies |\widehat{\theta}_n - \theta| \geq \epsilon \\ &\to 0 \qquad \text{since, } \widehat{\theta}_n \xrightarrow{P} \theta \end{aligned}$$

If $\theta_0 < \theta$, then similarly, for sufficiently large n, we have $(\theta_0 - \theta) + n^{-1/4} < -\epsilon < 0$, for some $\epsilon > 0$. Therefore,

$$
\begin{aligned}
P_\theta(Z_n \neq 0) &\leq P_\theta\left((\widehat{\theta}_n - \theta) \leq (\theta_0 - \theta) + n^{-1/4}\right) \\
&\leq P_\theta\left((\widehat{\theta}_n - \theta) \leq -\epsilon\right) \\
&\qquad \text{for sufficently large } n \text{ as,} \\
&\qquad (\widehat{\theta}_n - \theta) \leq ((\theta_0 - \theta) + n^{-1/4}) \implies (\widehat{\theta}_n - \theta) \leq (-\epsilon) \\
&\leq P_\theta\left(|\widehat{\theta}_n - \theta| \geq \epsilon\right) \\
&\qquad \text{since, } (\widehat{\theta}_n - \theta) \leq (-\epsilon) \implies |\widehat{\theta}_n - \theta| \geq \epsilon \\
&\to 0 \qquad \text{since, } \widehat{\theta}_n \xrightarrow{P} \theta
\end{aligned}
$$

Hence we have, $P_\theta(Z_n \neq 0) \to 0$, and hence, by definition of convergence in probability, $Z_n \xrightarrow{P} 0$ under θ.

Now, we know that, $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{d} N(0, \mathcal{I}^{-1}(\theta))$, and $Z_n = \sqrt{n}(T_n - \widehat{\theta}_n) \xrightarrow{P} 0$. Therefore, an application of Slutsky's theorem yields,

$$
\sqrt{n}(T_n - \theta) = \sqrt{n}(\widehat{\theta}_n - \theta) + Z_n \xrightarrow{d} N(0, \mathcal{I}^{-1}(\theta)) \tag{3}
$$

Case 2: Now we proceed to the case where $\theta = \theta_0$. In this case, consider the random variable,

$$
\sqrt{n}(T_n - \theta_0) = \begin{cases} \sqrt{n}(\widehat{\theta}_n - \theta_0) & \text{if } |\widehat{\theta}_n - \theta_0| > n^{-1/4} \\ 0 & \text{if } |\widehat{\theta}_n - \theta_0| \leq n^{-1/4} \end{cases}
$$

Similar to before, consider the following probability under $\theta = \theta_0$;

$$
\begin{aligned}
P_{\theta_0}\left(\sqrt{n}(T_n - \theta_0) \neq 0\right) &\leq P_{\theta_0}\left(|\widehat{\theta}_n - \theta_0| > n^{-1/4}\right) \\
&= P_{\theta_0}\left(|\sqrt{n}(\widehat{\theta}_n - \theta_0)| > n^{1/4}\right) \\
&\leq \frac{\mathrm{Var}_{\theta_0}(\widehat{\theta}_n)}{\sqrt{n}} \quad \text{, by Chebyshev's inequality} \\
&= \frac{\mathcal{I}^{-1}(\theta_0)}{\sqrt{n}} \\
&\xrightarrow{n\to\infty} 0
\end{aligned}
$$

Hence, if $\theta = \theta_0$, then under $\theta = \theta_0$, $\sqrt{n}(T_n - \theta_0) \xrightarrow{P} 0$, and as in probability convergence implies convergence in distribution, we have;

$$
\sqrt{n}(T_n - \theta_0) \xrightarrow{d} \delta_0 \tag{4}
$$

where δ_x is the distribution degenerate at x.

Thus, combining the equations (3) and (4), we obtain that under any $\theta \in \Theta$, the quantity $\sqrt{n}(T_n - \theta)$ follows an asymptotic normal distribution with mean 0 and variance $v(\theta)$, where $v(\theta)$ is given by;

$$v(\theta) = \begin{cases} \mathcal{I}^{-1}(\theta) & \text{if } \theta \neq \theta_0 \\ 0 & \text{if } \theta = \theta_0 \end{cases}$$

where the normal distribution with mean 0 and variance 0 is another notion of specifying the distribution degenerate at 0.

□

Thank you