$\forall \theta \in \Omega$ is a consistent sequence of estimators.

Theorem 10.1.6

$X_1, \dots, X_n$ iid. Let $\hat{\theta}$ be the MLE of θ. Let $\tau(\theta)$ be a continuous function of θ. Under certain regularity conditions on $f(x|\theta)$ (10.6.2, A1-A4 p. 516) we have $\forall \varepsilon > 0$ and every $\theta \in \Omega$ that $\lim_{n\to\infty} P(|\tau(\hat{\theta}) - \tau(\theta)| \geq \varepsilon) = 0$

i.e $\tau(\hat{\theta})$ is a consistent estimator for $\tau(\theta)$

Definition 10.1.7 and 10.1.9

For an estimator T_n, if $\lim_{n\to\infty} k_n \operatorname{Var}(T_n) = \tau^2 < \infty$, τ^2 is called a limiting variance. This is a hopeless definition. Some restriction on k_n is needed

If $k_n(T_n - \tau(\theta)) \xrightarrow{D} N(0, \sigma^2)$, σ^2 is called the asymptotic variance.

Example. Random sample $X_1, \dots, X_n$, $E[X_i] = \mu$, $0 < \operatorname{Var} X_i < \infty$, $i = 1, 2, \dots, n$

Then MLE for $\mu = \bar{X}_n$ and from CLT $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{D} N(0, \sigma^2)$

Also $\frac{1}{\bar{X}_n}$ is the MLE of $\frac{1}{\mu}$ and from the Delta theorem

$\sqrt{n}\left(\frac{1}{\bar{X}_n} - \frac{1}{\mu}\right) \xrightarrow{D} N\left(0, \frac{1}{\mu^4}\sigma^2\right)$ which shows that the asymptotic variance of $T_n = \frac{1}{\bar{X}_n}$ exist, $\mu \neq 0$

However, if $X_1, \dots, X_n$ are assumed to be normally distributed, the exact variance of $\frac{1}{\bar{X}_n} = \infty$, $\forall n$ and the limiting variance does not exist.

Example 10.1.10

A mixture distribution

$$Y_m / W_m \sim N(0, \; W_m + (1-W_m)\sigma_m^2)$$

$$W_m \sim B(p_m)$$

$$Var[Y_m] = E[Var[Y_m | W_m]] + Var[E[Y_m | W_m]]$$

$$= E[W_m + (1-W_m)\sigma_m^2] + Var[0]$$

$$= p_m + (1-p_m)\sigma_m^2$$

Let $p_m = 1 - \frac{1}{m}$ and $\sigma_m^2 = m^2$ $\Rightarrow$

$$Var[Y_m] = 1 - \frac{1}{m} + m \rightarrow \infty \quad \text{as } m \rightarrow \infty$$

Does $Y_m \xrightarrow{D} N(0,1)$?

But

$$P(Y_m \leq a) = P(Y_m \leq a \cap W_m = 1) + P(Y_m \leq A \cap W_m = 0)$$

$$= \underbrace{P(Y_m \leq a \mid W_m = 1)}^{N(0,1)} \cdot P(W_m = 1) + \overbrace{P(Y_m \leq a \mid W_m = 0)}^{N(0,\sigma_m^2)} \cdot P(W_m = 0)$$

$$= P(Z \leq a) \cdot p_m + \underbrace{P\left(Z \leq \frac{a}{\sigma_m}\right)}_{<1} \underbrace{(1 - p_m)}_{\downarrow 0}$$

$$\xrightarrow[m \to \infty]{} P(Z \leq a)$$

Hence $\lim Y_m \xrightarrow{D} N(0,1)$ where $\ell_m = 1, \; \forall m$

$\Rightarrow$ asymptotic variance is 1.

Definition 10.1.11

A sequence of estimators W_n is asymptotic efficient for a parameter $\tau(\theta)$ if $\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{D} N(0, v(\theta))$

and
$$v(\theta) = \frac{|\tau'(\theta)|^2}{E\left[\left(\frac{d}{d\theta}\log f(X|\theta)\right)^2\right]}$$

10.1.16

$\sqrt{n}(W_n - \tau(\theta)) \to N(0, \sigma_W^2)$

$\sqrt{n}(V_n - \tau(\theta)) \to N(0, \sigma_V^2)$

$\text{ARE}(V_n, W_n) = \dfrac{\sigma_W^2}{\sigma_V^2}$.

asymptotic relative efficiency of V_n w.r.t W_n

Theorem 10.1.2

$X_1, \dots, X_n$ iid $f(x|\theta)$. Let $\hat\theta$ be the MLE of θ and let $\tau(\theta)$ be a continuous function of θ. Under some regularity conditions on $f(x|\theta)$ (A1 - A6 p 516)

$$\sqrt{n}\left[\tau(\hat\theta) - \tau(\theta)\right] \xrightarrow{D} N(0, v(\theta))$$

where $v(\theta)$ is the Cramer-Rao lower bound.

Proof. Let us consider $\tau(\theta) = \theta$. Then

$$l(\theta|\underline{x}) = \log L(\theta|\underline{x}) = \log\left(\prod_{i=1}^n f(x_i|\theta)\right) = \sum_{i=1}^n \log f(x_i|\theta)$$

$$l'(\hat\theta_{n,e}|\underline{x}) = 0 \Rightarrow \sum_{i=1}^n \frac{d}{d\theta}\log f(x_i|\hat\theta_{n,e}) = 0 = \sum_{i=1}^n h(x_i|\hat\theta_{n,e})$$

Now $$h(x_i|\hat\theta_{n,e}) = h(x_i|\theta) + (\hat\theta_{n,e} - \theta)\left[\frac{d}{d\theta}h(x_i|\theta) + R(x_i|\hat\theta_{n,e}, \theta)\right]$$

$$\frac{d}{d\theta}h(x_i|\theta) = k(x_i|\theta) \qquad R(x_i|\hat\theta_{n,e}, \theta) = \cancel{(\hat\theta_{n,e} - \theta)}\left[c_1(\hat\theta_{n,e} - \theta) + c_2(\hat\theta_{n,e} - \theta)^2 + \dots\right]$$

Thereby
$$0 = \sum_{i=1}^n \frac{h(x_i|\hat\theta_{n,e})}{n} = \underbrace{\frac{\sum_{i=1}^n h(x_i|\theta)}{n}}_{\bar u_n} + (\hat\theta_{n,e} - \theta)\left(\underbrace{\frac{\sum_{i=1}^n k(x_i|\theta)}{n}}_{\bar v_n} + \underbrace{\frac{\sum_{i=1}^n R(x_i|\hat\theta_{n,e}, \theta)}{n}}_{\bar r_n}\right)$$

$$\Rightarrow \hat\theta_{n,e} - \theta = \frac{-\bar u_n}{\bar v_n + \bar r_n}$$

Substitute $\underline{x}$ by $\underline{X}$ gives.

$$\hat{\theta}_m - \theta = \frac{-\bar{U}_m}{\bar{V}_m + \bar{R}_m}$$

Now $U_i = S(X_i|\theta)$ are iid. $E[U_i]=0$, $Var[U_i] = I(\theta)$

$\Rightarrow \sqrt{m}\,\bar{U}_m \xrightarrow{D} N(0, I(\theta))$ (Central limit theorem 5.5.15)

V_i, $i = 1, 2, \dots, m$ are iid. $E[V_i] = -I(\theta)$ with finite variance $\Rightarrow \bar{V}_m \xrightarrow{P} -I(\theta)$ (weak law of large numbers 5.5.2)

and $\bar{R}_m \xrightarrow{P} 0$ since $\hat{\theta}_m - \theta \xrightarrow{P} 0$

$$\Rightarrow \sqrt{m}(\hat{\theta}_m - \theta) \xrightarrow{D} \frac{N(0, I(\theta))}{-I(\theta)} \sim N(0, v(\theta))$$

<u>Theorem 10.3.1</u>

$X_1, \dots, X_m$ iid ~~and~~, $\hat{\theta}$ is the MLE of θ and $f(x|\theta)$ satisfies the regularity conditions A1-A6 p. 516.

For a test $H_0: \theta = \theta_0$ vs. $H_1: \theta \neq \theta_0$, $-2\log\lambda(\underline{X})$ asymptotically has a $\chi^2(1)$ distribution i.e

$$-2\log\lambda(\underline{X}) \xrightarrow{D} \chi^2(1)$$

<u>Proof</u>. Let $\ell(\theta|\underline{x}) = \log L(\theta|\underline{x})$. A Taylor expansion around $\hat{\theta}_{m,e}$ gives:

$$\ell(\theta|\underline{x}) = \ell(\hat{\theta}_{m,e}|\underline{x}) + \underbrace{\ell'(\hat{\theta}_{m,e}|\underline{x})}_{=0}(\theta - \hat{\theta}_{m,e}) + \frac{\ell''(\hat{\theta}_{m,e}|\underline{x})(\theta - \hat{\theta}_{m,e})^2}{2} + \dots$$

such that $l(\theta_0|\underline{x}) = l(\hat{\theta}_{m,e}|\underline{x}) + \dfrac{l''(\hat{\theta}_{m,e}|\underline{x})(\theta_0 - \hat{\theta}_{m,e})^2}{2} + \dots$

and $-2\log\lambda(\underline{x}) = -2\log\left[\dfrac{L(\theta_0|\underline{x})}{L(\hat{\theta}_{m,e}|\underline{x})}\right] = -2\,l(\theta_0|\underline{x}) + 2\,l(\hat{\theta}_{m,e}|\underline{x})$

$$= -\,l''(\hat{\theta}_{m,e}|\underline{x})(\theta_0 - \hat{\theta}_{m,e})^2 + R_m$$

If $\theta = \theta_0$, we have

$$\sqrt{m}(\hat{\theta}_m - \theta_0) \xrightarrow{D} N(0, I(\theta_0)^{-1})$$

and substituting $\underline{X}$ for $\underline{x}$ we get

$$-\frac{l''(\hat{\theta}_m|\underline{X})}{m} = -\sum_{i=1}^{m}\frac{l''(\hat{\theta}_m|X_i)}{m}$$

We have $\displaystyle\sum_{i=1}^{m}\frac{l''(\theta_0|X_i)}{m} = \bar{V}_m \xrightarrow{P} -I(\theta_0)$

$\hat{\theta}_m \xrightarrow{P} \theta_0 \Rightarrow \quad l''(\hat{\theta}_m|X_i) \xrightarrow{P} l''(\theta_0|X_i)$ (continuous).

Therefore $\displaystyle -\sum_{i=1}^{m}\frac{l''(\hat{\theta}_m|X_i)}{m} \xrightarrow{P} I(\theta_0)$

Hence $\displaystyle -2\log\lambda(\underline{X}) = -\frac{l''(\hat{\theta}_m|\underline{X})(\sqrt{m}(\hat{\theta}_m - \theta_0))^2}{m} + \underset{\xrightarrow{P} 0}{R_m}$

$$\xrightarrow{D} I(\theta_0)[N(0, I(\theta_0)^{-1})]^2$$

If $u \sim N(0, I(\theta_0)^{-1}) \Rightarrow Z = \sqrt{I(\theta_0)}\,u \sim N(0,1)$

and $Z^2 \sim I(\theta_0)u^2 \sim \chi^2(1)$.

Theorem 10.3.3.