b) If $\lim_{n\to\infty} V_\theta(T_n) = 0$ and

$\lim_{n\to\infty} E_\theta(T_n) = \tau(\theta)$ $\forall \theta \in \Theta$,

then T_n is a consistent estimator of $\tau(\theta)$.

proof] a) Using Gen Cheb Ineq with $Y = T_n$, $u(T_n) = (T_n - \tau(\theta))^2$ and $c = \varepsilon^2$ shows that for any $\varepsilon > 0$,

$$P_\theta[|T_n - \tau(\theta)| \geq \varepsilon] =$$

$$P_\theta[(T_n - \tau(\theta))^2 \geq \varepsilon^2] \leq \frac{\overbrace{E_\theta[(T_n - \tau(\theta))^2]}^{\to 0 \text{ as } n\to\infty}}{\varepsilon^2} \to 0.$$

b) Since $MSE_{\tau(\theta)}(T_n) = V_\theta(T_n) + [E_\theta(T_n) - \tau(\theta)]^2$, $MSE_{\tau(\theta)}(T_n) \to 0$

if both $V_\theta(T_n) \to 0$ and $E_\theta(T_n) \to \tau(\theta)$, the result follows by a).

31] see p49. a) Let $r > 0$. Y_n converges in rth mean to Y, written $Y_n \xrightarrow{r} Y$, if $E[|Y_n - Y|^r] \to 0$ as $n \to \infty$.

$Y_n \xrightarrow{2} Y$ is also known as "converges in quadratic mean" or mean square convergence.

b) $Y_n \xrightarrow{r} \tau(\theta)$ if

$$E_\theta[|Y_n - \tau(\theta)|^r] \to 0 \text{ as } n \to \infty.$$

c) $MSE_{\tau(\theta)}(T_n) \to 0$ iff $Y_n \xrightarrow{2} \tau(\theta)$.

Note: a) $|Y_n - Y|^2 = (Y_n - Y)^2$.

b) If $r > 1$, $Y_n \xrightarrow{r} Y$ is often written as

$Y_n \xrightarrow{L^r} Y$ or $Y_n \xrightarrow{L^r} Y$.

32} See p. 576. a) X_n <u>converges with probability one</u> to X, written $X_n \xrightarrow{wp1} X$, if $P\left(\lim_{n\to\infty} X_n = X\right) = 1$,

b) $X_n \xrightarrow{wp1} \tau(\theta)$ if $P\left(\lim_{n\to\infty} X_n = \tau(\theta)\right) = 1$.

c) This convergence is also known as i) <u>strong convergence</u>

ii) <u>converges almost surely</u> $X_n \xrightarrow{as} X$

iii) <u>converges almost everywhere</u>

$X_n \xrightarrow{ae} X$.

33} Let Y_n be a sequence of iid RVs with $E(Y_n) = \mu$.

a) Strong Law of Large Numbers <u>SLLN</u>:

$\bar{X}_n \xrightarrow{wp1} \mu.$

b) Weak Law of Large Numbers WLLN:

$\bar{X}_n \xrightarrow{P} \mu.$

proof of WLLN if $V(Y_n) = \sigma^2$:

By Chebyshev's Inequality, for any $\varepsilon > 0$,

$$P(|\bar{Y}_n - \mu| \geq \varepsilon) = P[(\bar{Y}_n - \mu)^2 \geq \varepsilon^2]$$

$V(\bar{Y}_n) = E[(\bar{Y}_n - \mu)^2]$

$$\leq \frac{V(\bar{Y}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ as } n \to \infty.$$

$\square$

Warning: Students often forget that $V(Y_n) = \sigma^2$ is not needed. $\bar{X}_n \xrightarrow{wp1} \mu$ and $\bar{X}_n \xrightarrow{P} \mu$ if the X_i are iid with $E(X_i) = \mu$.

LS 15

34} Let $k > 0$. If $E(X^k)$ is finite, then $E(X^j)$ is finite for $0 < j \le k$.

Proof: If $|y| \le 1$, then $|y^j| = |y|^j \le 1$.

If $|y| > 1$, then $|y|^j \le |y|^k$.

Hence $|y|^j \le |y|^k + 1$ and

$|X|^j \le |X|^k + 1$.

$\therefore \quad E\left[|X|^j\right] \le E\left[|X|^k\right] + 1 < \infty$. $\square$

Note: moments use j, k positive integers.
$E[Y] < \infty$ does not mean $E(Y^{-1})$ exists.

35} Jensen's Inequality:

$g[E(X)] \le E[g(X)]$ if the expected values exist, and g is convex on an interval containing the range of X.

Note: a) Let (a,b) be an open interval where $a=-\infty$ and $b=\infty$ are allowed. A sufficient condition for a function g to be convex on an open interval (a,b) is $g''(x)>0$ on (a,b).

If $(a,b)=(0,\infty)$ and g is continuous on $[0,\infty)$ and convex on $(0,\infty)$, then g is convex on $[0,\infty)$.

b) If X is a positive RV then the range of X is $(0,\infty)$.

36) If $X_n \xrightarrow{r} X$, then

$X_n \xrightarrow{k} X$ where $0<k<r$.

Proof) Let $U_n = |X_n - X|^r$ and

$W_n = |X_n - X|^k$. Then

$U_n = W_n^{\,t}$ where $t = \frac{r}{k} > 1$

and $g(x) = x^t$ is convex on $[0, \infty)$.

$(g'(x) = t x^{t-1},\ g''(x) = t(t-1) x^{t-2})$

By Jensen's Inequality,

$$E[|X_n - X|^r] = E[U_n] = E[W_n^{\,t}]$$

$$\ge (E[W_n])^t = (E[|X_n - X|^k])^{\frac{r}{k}}, \quad r > k.$$

$\therefore \lim_{n \to \infty} E[|X_n - X|^r] = 0 \Rightarrow$

$\lim_{n \to \infty} E[|X_n - X|^k] = 0, \quad 0 < k < r.$

373* If $X_n \xrightarrow{r} X$ then $X_n \xrightarrow{P} X$.

16.5

proof) $|X_n - X|^r \geq |X_n - X|^r \underbrace{I[|X_n - X| \geq \varepsilon]}_{0 \text{ or } 1}$

$\geq \varepsilon^r I[|X_n - X| \geq \varepsilon]$,

So for any $\varepsilon > 0$,

$E[|X_n - X|^r] \geq E[|X_n - X|^r I(|X_n - X| \geq \varepsilon)]$

$\geq E[\varepsilon^r I(|X_n - X| \geq \varepsilon)]$

$= \varepsilon^r P[|X_n - X| \geq \varepsilon]$.

$\therefore P[|X_n - X| \geq \varepsilon] \leq \dfrac{E[|X_n - X|^r]}{\varepsilon^r} \to 0$

as $n \to \infty$.

$\square$

38}* a) If $X_n \xrightarrow{wp1} X$, then $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{D} X$.

p71 b) If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$.

c) If $X_n \xrightarrow{r} X$, then $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{D} X$.

p66 d) $X_n \xrightarrow{P} \tau(\theta)$ iff $X_n \xrightarrow{D} \tau(\theta)$.

e) If $X_n \xrightarrow{D} X$ and $X_n \xrightarrow{D} Y$ then $X \overset{D}{=} Y$ and $F_X(x) = F_Y(x) \quad \forall\, x \in \mathbb{R}$.

we will prove some of this later.

For c) see 37}.

proof of e) Suppose X has cdf F and Y has cdf G.

Then F and G agree at their common points of continuity. Hence at all but countably many points since F and G are cdfs.

Hence at all points by right continuity. □

Note: a) if $X_n \xrightarrow{a} X$ and $X_n \xrightarrow{a} Y$,

then $X \stackrel{D}{=} Y$ since $\xrightarrow{a} \Rightarrow \xrightarrow{D}$

where $a =$ wpl, r, or P.

b)

$$\begin{tikzcd} \boxed{\xrightarrow{r}} \arrow[dr, Rightarrow] & & \\ & \boxed{\xrightarrow{P}} \arrow[r, Rightarrow] & \boxed{\xrightarrow{D}} \\ \boxed{\xrightarrow{wpl}} \arrow[ur, Rightarrow] & & \end{tikzcd}$$

39) $A \Rightarrow B$ does not mean that if A does not hold, then B does not hold. $A \Rightarrow B$ means that if A holds, then B holds.

ex) Let $P(X_n = n) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$

Then $E(X_n) = n \frac{1}{n} = 1 \quad \forall n.$

x	0	n
$P(X_n = x)$	$1-\frac{1}{n}$	$\frac{1}{n}$

$\therefore E[|X_n - 0|] = E[X_n] = 1 \quad \forall n$ and X_n

does not satisfy $X_n \xrightarrow{1} 0$.

Let $\varepsilon > 0$.

$$P[|X_n - 0| \geq \varepsilon] \leq P(X_n = n) = \frac{1}{n} \to 0$$

equality once $n > \varepsilon$, so for $0 < \varepsilon < 1$

as $n \to \infty$.

$\therefore X_n \xrightarrow{P} 0 \quad \therefore X_n \xrightarrow{D} 0.$

point mass at 0

ex) Let $P(X_n = 0) = 1 - \frac{1}{n}$ and $P(X_n = 1) = \frac{1}{n}$.

Then $E[(X_n - 0)^2] = E[X_n^2] = \sum x^2 P(X_n = x)$

x	0	1
$P(X_n = x)$	$1-\frac{1}{n}$	$\frac{1}{n}$

$$= 0^2\left(1 - \frac{1}{n}\right) + 1^2 \frac{1}{n} = \frac{1}{n} \to 0$$

$\therefore X_n \xrightarrow{2} 0$ and $X_n \xrightarrow{P} 0$ and $X_n \xrightarrow{D} 0$.

Note that $E[|X_n - 0|] = E[X_n] = \frac{1}{n} \to 0$.

So $X_n \xrightarrow{1} 0$ as expected since $X_n \xrightarrow{2} 0$.

40} Know <u>Standard Limit Theorem</u>

Let $\hat{\theta}_n$ be the MLE or UMVUE of θ. If $\tau'(\theta) \neq 0$, then under strong regularity conditions,

$$\sqrt{n}\left[\tau(\hat{\theta}_n) - \tau(\theta)\right] \xrightarrow{D} N\left[0, \frac{[\tau'(\theta)]^2}{I_1(\theta)}\right].$$

p459

41} If $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N\left[0, \frac{1}{I_1(\theta)}\right]$,

then 40} holds by the Delta Method.

42} If 1st and 2nd derivatives can be interchanged with the integral

$$\frac{d}{d\theta}\int f(y|\theta)\,dy = \int \frac{d}{d\theta} f(y|\theta)\,dy$$

and $\frac{d^2}{d\theta^2}\int f(y|\theta)\,dy = \int \frac{d^2}{d\theta^2} f(y|\theta)\,dy$

then ii) and iii) hold.

See HW3 #4.

Hence ii) and iii) hold for a 1P-REF,

If i) the Y_i are iid with pdf $f(y|\theta)$ and likelihood function $L(\theta) = \prod_{i=1}^n f(y_i|\theta)$, ii) $E_\theta\left[\frac{d \log L(\theta)}{d\theta}\right] = 0$

and iii) $E_\theta\left[\left(\frac{d \log L(\theta)}{d\theta}\right)^2\right] = -E_\theta\left[\frac{d^2 \log L(\theta)}{d\theta^2}\right]$

exists and is nonzero for all θ in a neighborhood of the true value θ_0, then for the MLE,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\left(0, \frac{1}{I_1(\theta_0)}\right). \quad (*)$$

43) Hence $(*)$ holds for a 1P-REF with a pdf.

44) A 1P-REF has pdf or pmf

$f(y) = f(y|\theta) = h(y)\, c(\theta) \exp[w(\theta)\, t(y)]$

Some other conditions also need to hold.

1P-REF $\qquad I_1(\theta)$

$Y \sim bin(k, \rho)$ k known $\qquad I_1(\rho) = \frac{k}{\rho(1-\rho)}$

$Y \sim EXP(\lambda)$ $\qquad I_1(\lambda) = \frac{1}{\lambda^2}$

$Y \sim G(\nu, \lambda)$ ν known $\qquad I_1(\lambda) = \frac{\nu}{\lambda^2}$

$Y \sim N(\mu, \sigma^2)$ $\begin{cases} \sigma \text{ known} & I_1(\mu) = \frac{1}{\sigma^2} \\ \mu \text{ known} & I_1(\sigma^2) = \frac{1}{2\sigma^4} \end{cases}$

$Y \sim Pois(\theta)$ $\qquad I_1(\theta) = \frac{1}{\theta}$,

See exam 1 review,

$\sqrt{n}(\overline{Y} - \theta) \xrightarrow{D} N(0, \tau^2)$ then $\tau^2 = \frac{1}{I_1(\theta)}$

45) P.464-8 203

If $Y_1, ..., Y_n$ are iid from a 1P-REF

then a) $I_n(\theta) = n I_1(\theta)$.

b) $$I_n[\tau(\theta)] = \frac{n I_1(\theta)}{[\tau'(\theta)]^2} \text{ if } \tau'(\theta) \neq 0,$$

So $$I_1(\tau(\theta)) = \frac{I_1(\theta)}{[\tau'(\theta)]^2}.$$

c) $$I_1(\theta) = -E_\theta\left[\frac{d^2 \log[f(Y|\theta)]}{d\theta^2}\right].$$

46) For a 1P-REF $\dfrac{\sum_{i=1}^n t(Y_i)}{n}$ is

the UMVUE of $\mu_T = E[t(Y_i)]$

and often the MLE of μ_T.

Let $V(t(Y_i)) = \sigma_t^2$.

By the CLT,

$$\sqrt{n}\left[\frac{\sum_{i=1}^{n} t(Y_i)}{n} - \mu_t\right] \xrightarrow{D} N(0, \sigma_t^2).$$

Take $W_i = t(Y_i)$. Then $\frac{\sum_{i=1}^{n} t(Y_i)}{n} = \overline{W}_n$

with $E(W_i) = \mu_t$ and $V(W_i) = \sigma_t^2$.

If $\mu_t = \tau(\theta)$, then often $\sigma_t^2 = \frac{[\tau'(\theta)]^2}{I_1(\theta)}$, and $\overline{W}_n = \tau(\hat{\theta})$,

by the Standard Limit Theorem.

47) Let $Y_1, \ldots, Y_n$ be the data.

Let $\overset{\text{min}}{Y_{(1)}} \le Y_{(2)} \le \cdots \le Y_{(n-1)} \le \underset{\uparrow \text{ max}}{Y_{(n)}}$

be the <u>order statistics</u>

The notation $Y_{i:n} = Y_{(i)}$ is also used.

The sample median

$$MED(n) = \begin{cases} Y_{\left(\frac{n+1}{2}\right)} & n \text{ odd} \\ \dfrac{Y_{\left(\frac{n}{2}\right)} + Y_{\left(\frac{n}{2}+1\right)}}{2} & n \text{ even} \end{cases}$$

MED(n) is the "middle order statistic" or the "average of the 2 middle order statistics."

ex) $Y_1 = 1$, $Y_2 = 4$, $Y_3 = 2$, $Y_4 = 5$, $Y_5 = 3$

1	2	3	4	5
$Y_{(1)}$	$Y_{(2)}$	$Y_{(3)}$	$Y_{(4)}$	$Y_{(5)}$

$MED(n) = Y_{\left(\frac{n+1}{2}\right)} = Y_{(3)} = 3$

48) The population median MED(Y) is any value such that $P(Y \le MED(Y)) \ge \frac{1}{2}$ and $P(Y \le MED(Y)) \ge \frac{1}{2}$. MED(Y)

21.5

is unique if the pdf satisfies

$f(MED(Y)) > 0$.

If the pdf is symmetric about μ,

then μ is a pop median.

49) Thm Let $Y_1, \ldots, Y_n$ be iid with a pdf

(pql know)

that is positive at the (unique) pop

median: $f(MED(Y)) > 0$. Then

$$\sqrt{n}\,[MED(n) - MED(Y)] \xrightarrow{D} N\left[0, \frac{1}{4[f(MED(Y))]^2}\right].$$

Note: Estimating $f(MED(Y))$ is harder

than estimating $V(Y)$. Compare CLT.

50) If $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma_A^2)$

then $T_n \approx N\left(\theta, \frac{\sigma_A^2}{n}\right)$ written

$T_n \sim AN\left(\theta, \frac{\sigma_A^2}{n}\right)$ an approximate

distribution. The limiting

distribution $N(0, \sigma_A^2)$ does not depend on n, but the approximate distribution often depends on n.

ex) Y_i iid with mean μ variance σ^2

$$\bar{Y}_n \sim AN\left(\mu, \frac{S^2}{n}\right),$$

$$S_n^2 = S^2 = \text{sample variance} = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2.$$

S_M^2 = method of moments estimator of σ^2

$$= \frac{n-1}{n} S^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2.$$

51) Let $T_n = T_n(Y_1, \ldots, Y_n)$ be an estimator of parameter μ_T such that $\sqrt{n}(T_n - \mu_T) \xrightarrow{D} N(0, \sigma_A^2)$.

The <u>asymptotic variance</u> of T_n is $\frac{\sigma_A^2}{n}$. If S_A^2 is a consistent estimator of σ_A^2, then the (asymptotic) standard error (SE) of T_n is $SE(T_n) = \frac{S_A}{\sqrt{n}}$.

So $T_n \sim AN(\mu_T, \frac{S_A^2}{n})$.

If $Y_1, ..., Y_n$ are iid with cdf F then $\sigma_A^2 = \sigma_A^2(F)$ depends on F,

Note: p232 often $n V(T_n) \xrightarrow{P} \sigma_A^2$, but not always.

<u>end Exam I material</u>

5.23 know p239 Let T_{1n} and T_{2n} be two estimators of a parameter θ