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

and natural parameterization $h(y)\, b(\eta) \exp[\eta\, t(y)]$

where $\eta = w(\theta)$.

see Geisser's result from large sample theory

Let $E(t(Y)) = \mu_t = g(\eta)$ and

$V(t(Y)) = \sigma_t^2$. Then $I_1(\eta) = \sigma_t^2 = g'(\eta) = \dfrac{[g'(\eta)]^2}{I_1(\eta)}$

and $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} t(Y_i) - \mu_t\right) \xrightarrow{D} \underbrace{N(0, \sigma_t^2)}_{CLT} = N\left[0, \dfrac{[g'(\eta)]^2}{I_1(\eta)}\right]$

§8.2 8] p 223 (know for qual) Let $Y_1, \ldots, Y_n$ be iid with cdf F.

Let T_{1n} and $T_{2n} \equiv T_{2n}(\underline{Y})$ be estimators of θ such that $n^\delta (T_{1n} - \theta) \xrightarrow{D} N(0, \sigma_1^2(F))$

and $n^\delta (T_{2n} - \theta) \xrightarrow{D} N(0, \sigma_2^2(F))$.

the better estimator has the smaller $\sigma_i(F)$

Then the asymptotic relative efficiency of T_{1n} with respect to T_{2n} is

$$ARE(T_{1n}, T_{2n}) = \frac{\sigma_2^2(F)}{\sigma_1^2(F)}. \quad \text{Typically } \delta = 0.5.$$

T_{1n} is "better" than T_{2n} if $ARE(T_{1n}, T_{2n}) > 1$

T_{1n} is "worse" if $ARE(T_{1n}, T_{2n}) < 1$.

9] p 225 know for qual. Let $\tau'(\theta) \neq 0$. An estimator T_n of $\tau(\theta)$ is asymptotically efficient

if $\sqrt{n}(T_n - \tau(\theta)) \xrightarrow{D} N\left(0, \dfrac{[\tau'(\theta)]^2}{I_1(\theta)}\right) = N(0, FCRLB_1(\tau(\theta)))$.

So T_n is an asymptotically efficient estimator of θ

if $\sqrt{n}(T_n - \theta) \xrightarrow{D} N\left(0, \dfrac{1}{I_1(\theta)}\right) = N(0, FCRLB_1(\theta))$

and $T_n \approx N\left(\theta, \frac{1}{nI_1(\theta)}\right) = N(\theta, FCRLB_n(\theta))$ $\quad T_n \sim AN\left(\theta, \frac{1}{nI_1(\theta)}\right)$.

10] *Rule of thumb 5] says MLE $\tau(\hat{\theta}_n)$ is an asymptotically efficient estimator of $\tau(\theta)$.
7] Says $\frac{1}{n}\sum t(Y_i)$ is an asymptotically efficient estimator of $E[t(Y)] = g(\eta) = \mu_t$

rule of thumb p225-6, 244

[Under strong regularity conditions, if U_n is the UMVUE of $\tau(\theta)$, then U_n is asymptotically efficient.

§8.3 11] Large sample methods give a useful approx to a statistic. In the CLT, the limiting distribution of $\sqrt{n}(\bar{Y} - \mu)$ is $N(0, \sigma^2)$, and $\bar{Y} \approx N(\mu, \frac{\sigma^2}{n})$.

To be useful, we need to know how large n needs to be for the approx to be used. For the CLT, if the pmf or pdf is not highly skewed, $n \geq 30$ often works.

12] Know p226 Let $\{Z_n, n = 1, 2, \ldots\}$ be a sequence of random variables with cdfs F_n and let X be a RV with cdf F. Z_n converges in distribution (or law) to X, written $Z_n \xrightarrow{D} X$, if

$$\lim_{n\to\infty} F_n(t) = F(t)$$ at each continuity point t of F. X is the limiting dist or asymptotic dist of Z_n. Tip! limiting

ex] Let $Z_n = \sqrt{n}\left(\frac{\overline{Y}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1)$ by CLT.

So $\lim_{n\to\infty} F_n(t) = \underbrace{\Phi(t)}_{N(0,1)\text{ cdf}} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,dy \quad \forall\, t \in \mathbb{R}$

is continuous

13] If $X_n \xrightarrow{D} X$, then $P(a < X_n \leq b) = F_n(b) - F_n(a)$ $\to F(b) - F(a) = P(a < X \leq b)$ if F is continuous at a and b. So F can be used to approximate probabilities and percentiles (for confidence intervals and tests of hypotheses $x_{n,1-\alpha} \approx x_{1-\alpha}$ where $P(X_n \leq x_{n,1-\alpha}) = 1-\alpha$).

Read ex 8.8 and 8.9 carefully.

14] * p 227 X_n converges in probability to X, written $X_n \xrightarrow{P} X$, if $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$ for all $\varepsilon > 0$.
$\iff \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1 \quad \forall\, \varepsilon > 0$.

15] Know for E3, final, qual p 267 A sequence of estimators T_n is consistent for $\tau(\theta)$ if $T_n \xrightarrow{P} \tau(\theta)$ $\forall\, \theta \in \Theta$. If T_n is consistent for $\tau(\theta)$, then T_n is a consistent estimator of $\tau(\theta)$.

16] Know p 229 – 230

a) If $\lim_{n\to\infty} MSE_{\tau(\theta)}(T_n) = 0$ $\forall\, \theta \in \Theta$, T_n is a consistent estimator of $\tau(\theta)$.

b) If $\lim_{n\to\infty} V_\theta(T_n) = 0 \ \forall \theta \in \Theta$ and $\lim_{n\to\infty} E_\theta(T_n) = \tau(\theta)$ $\forall \theta \in \Theta$, then T_n is a consistent estimator of $\tau(\theta)$

c) Let $0 < \delta \le 1$. If $n^\delta (T_n - \tau(\theta)) \xrightarrow{D} N(0, v(\theta))$ $\forall \theta \in \Theta$, then T_n is a consistent estimator of $\tau(\theta)$.

d) Weak law of large numbers (WLLN): $Y_1, \ldots, Y_n$ be iid with $E(Y_i) = \mu$, then $\overline{Y}_n \xrightarrow{P} \mu$.

ex] Let $Y_1, \ldots, Y_n$ be iid $f(y|\underline{\theta})$, $\underline{\theta} \in \Theta$, with $E_{\underline{\theta}}(Y_i) = \mu$ and $V_{\underline{\theta}}(Y_i) = \sigma^2$.

i) $MSE_\mu(\overline{Y}_n) = V_{\underline{\theta}}(\overline{Y}) + (E_\theta \overline{Y} - \mu)^2 = \frac{\sigma^2}{n} + 0 \to 0$

ii) $V_{\underline{\theta}}(\overline{Y}_n) = \frac{\sigma^2}{n} \to 0$, $E_{\underline{\theta}} \overline{Y} = \mu \to \mu$

iii) $\sqrt{n}(\overline{Y}_n - \mu) \xrightarrow{D} N(0, \sigma^2)$ by CLT

iv) $\overline{Y}_n \xrightarrow{P} \mu$ by WLLN

So i), ii), iii), or iv) show $\overline{Y}_n$ is a consistent estimator of $\mu = E_{\underline{\theta}}(Y_i)$.

17] p228 Generalized Chebyshev's Inequality: Let $U: \mathbb{R} \to [0, \infty)$ be a nonnegative function and $E[U(Y)]$ exist. Then for any $c > 0$

$$P[U(Y) \ge c] \le \frac{E[U(Y)]}{c}.$$

p228 proof $E[U(Y)] = \int_{\mathbb{R}} u(y) f(y)\,dy = \int_{\{y: u(y) \geq c\}} u(y) f(y)\,dy + \int_{\{y: u(y) < c\}} u(y) f(y)\,dy$

$\geq \int_{\{y: u(y) \geq c\}} u(y) f(y)\,dy \geq c \int_{\{y: u(y) \geq c\}} f(y)\,dy = c\,P[U(Y) \geq c]$.

$u(y)f(y) \geq 0$

Replace integrals by sums for pmfs.

18) p228 Chebyshev's inequality: If $V(Y)$ exists,

$$P[|Y - \mu| \geq c] \leq \frac{V(Y)}{c^2}.$$

proof Take $u(y) = (y - \mu)^2$. Then

$$P[|Y - \mu| \geq c] = P[(Y - \mu)^2 \geq c^2] \leq \frac{V(Y)}{c^2}.$$

p230 Proof of WLLN when $V(Y_i) = \sigma^2$:

For any $\varepsilon > 0$, by Cheb ineq,

$$P[|\overline{Y}_n - \mu| \geq \varepsilon] \leq \frac{V(\overline{Y}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0.$$

So $\overline{Y}_n \xrightarrow{P} \mu$.

p229 Proof that $MSE_{\tau(\theta)}(T_n) \to 0 \Rightarrow T_n$ is consistent for $\tau(\theta)$.

For any $\theta \in \Theta$ and any $\varepsilon > 0$, $P_\theta[|T_n - \tau(\theta)| \geq \varepsilon] = P_\theta[(T_n - \tau(\theta))^2 \geq \varepsilon^2]$

$$\leq \frac{E_\theta[(T_n - \tau(\theta))^2]}{\varepsilon^2} = \frac{MSE_{\tau(\theta)}(T_n)}{\varepsilon^2} \to 0.$$

So $T_n \xrightarrow{P} \tau(\theta) \quad \forall \; \theta \in \Theta$.

19] Any estimator T_n of $\tau(\theta)$ should be consistent, but we would like to know how large an n is needed for T_n to be close to $\tau(\theta)$ with high prob.

§8.4 20] p230 53.5

Slutsky's Theorem: Suppose $Y_n \xrightarrow{D} Y$ and $W_n \xrightarrow{P} w$ for some constant w.

a) $Y_n + W_n \xrightarrow{D} Y + w$

b) $Y_n W_n \xrightarrow{D} wY$

c) $Y_n / W_n \xrightarrow{D} Y/w$ if $w \neq 0$.

21] If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$.

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

22] p231 **Continuous Mapping Theorem**:

If $X_n \xrightarrow{D} X$ and g is continuous, $g(X_n) \xrightarrow{D} g(X)$.

ex) $\sqrt{n}\left(\frac{\bar{Y}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1)$, so

$$n\left(\frac{\bar{Y}_n - \mu}{\sigma}\right)^2 \xrightarrow{D} [N(0,1)]^2 \sim \chi^2_1.$$

23] p232 **Continuity Theorem**: Let Y_n have mgf m_n and Y have mgf m where m_n and m are defined for $|t| \le d$ for some $d > 0$. If $m_n(t) \to m(t)$ $\forall\, |t| < c$ where $0 < c < d$, then $Y_n \xrightarrow{D} Y$.

Skip §8.5 – 8.7.

580 54

could use θ_0

Second order delta method!

24] Not on final, sometimes on qual. p 245 Fix θ.

Th Suppose $\sqrt{n}(T_n - \theta) \xrightarrow{D} N[0, \tau^2(\theta)]$, $g'(\theta) = 0$ and $g''(\theta) \neq 0$. Then delta method does not apply, but

$$n\left[g(T_n) - g(\theta)\right] \xrightarrow{D} \frac{1}{2}\tau^2(\theta)\, g''(\theta)\, \chi_1^2.$$

ex) 8.26, 8.27c

ex 8.14] $X_n \sim \text{bin}(n, p)$ $0 < p < 1$, $g(\theta) = \theta^3 - \theta$.

Find limiting distribution of $n\left[g\left(\frac{X_n}{n}\right) - c\right]$

for appropriate constant c when $p = \frac{1}{\sqrt{3}}$.

Soln Let $\theta = p$. Then $g'(\theta) = 3\theta^2 - 1$ and $g''(\theta) = 6\theta$.

$g\left(\frac{1}{\sqrt{3}}\right) = \left(\frac{1}{\sqrt{3}}\right)^3 - \frac{1}{\sqrt{3}} = \frac{1}{\sqrt{3}}\left(\frac{1}{3} - 1\right) = \frac{-2}{3\sqrt{3}} = c$.

$g'\left(\frac{1}{\sqrt{3}}\right) = 0$, $g''\left(\frac{1}{\sqrt{3}}\right) = \frac{6}{\sqrt{3}}$.

$\sqrt{n}\left(\frac{X_n}{n} - p\right) \xrightarrow{D} N(0, p(1-p))$ by CLT

so $\tau^2(p) = p(1-p)$.

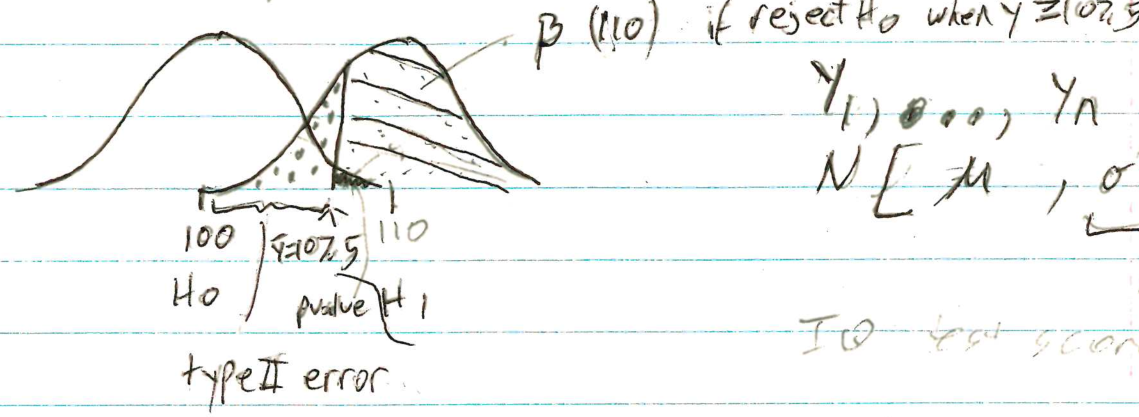
$\tau^2\left(\frac{1}{\sqrt{3}}\right) = \frac{1}{\sqrt{3}}\left(1 - \frac{1}{\sqrt{3}}\right)$. So

$$n\left[g\left(\frac{X_n}{n}\right) - \frac{-2}{3\sqrt{3}}\right] \xrightarrow{D} \frac{1}{2}\frac{1}{\sqrt{3}}\left(1 - \frac{1}{\sqrt{3}}\right)\frac{6}{\sqrt{3}}\chi_1^2$$

$$= \left(1 - \frac{1}{\sqrt{3}}\right)\chi_1^2.$$

More on power

ex



$Y_1, \ldots, Y_n$ iid $N[\mu, \underbrace{\sigma^2 = (15)^2}_{\text{known}}]$

IQ test scores

PDQ statistics Norman & Streiner p70-73

SRS of $n = 25$ $H_0\ \mu = 100$ $H_A\ \mu = 110$

observe $\bar{y} = 107.5$ pvalue =

P_{H_0}[test statistic ($\bar{Y}$) is "at least as extreme" as the statistic actually observed ($\bar{y}$)] when H_0 is true where "at least as extreme" gives pvalue $= P(\bar{Y} \geq 107.5)$ since this is a "right tail test" with $\mu_1 = 110 > \mu_0 = 100$.

see 7.20C

Suppose we reject H_0 if $\bar{Y} \geq 107.5$,

then pvalue $= P_{100}(\text{reject } H_0) = \underset{\mu=100}{\beta(100)} = P_{\mu=100}(\bar{Y} \geq 107.5)$

$$= P_{\mu=100}\left(\frac{\bar{Y} - 100}{15/\sqrt{25}} \geq \frac{107.5 - 100}{15/\sqrt{25}}\right) = P\left(Z > \frac{7.5}{3}\right)$$

$\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$= P(Z > 2.50) = 1 - .9938 = 0.0062$$

$Z \sim N(0,1)$

	.00
2.5	.9938

Let the rejection region be $(\bar{Y} \geq 107.5)$.

So P(type I error) $= \alpha =$ pvalue $= 0.0062$. If H_1 is true, power $= \beta(110) = P_{\mu=110}(\bar{Y} > 107.5) = P\left(\frac{\bar{Y} - 110}{15/\sqrt{25}} > \frac{107.5 - 110}{15/\sqrt{25}}\right) = P(Z > -0.833)$

$\approx 1 - P(Z < -0.83) = 1 - .2033 = \boxed{0.7967}$

	.03
-0.8	.2033

= power when $\mu = 110$.

$P(\text{Type II error}) = P_{110}(\text{fail to reject } H_0)$

$= 1 - .7967 = 0.2033$ if $\bar{Y} = 107.5$.

ex] $H_0\ \mu = 100$ $\quad H_A\ \mu = 110$ $\quad$ reject H_0 if $\bar{Y} \geq 105$

$\bar{y} = 107.5$ was observed $\quad Y_1, \ldots, Y_n$ iid $N(\mu, \sigma^2 = (15)^2)$

$n = 25$

Then $\text{pvalue} = P_{100}(\bar{Y} \geq \bar{y}) = P(Z \geq 2.50) = .0062$

as before.

$\alpha = P_{100}(\bar{Y} \geq 105) = P\left(\frac{\bar{Y} - 100}{15/\sqrt{25}} > \frac{105 - 100}{15/\sqrt{25}}\right)$

$= P(Z > \frac{5}{3}) \approx P(Z > 1.67) = 1 - .9525 = .0475$

$\beta(110) = P_{110}(\text{reject } H_0) = P_{110}(\bar{Y} \geq 105) =$

$P\left(\frac{\bar{Y} - 110}{15/\sqrt{25}} > \frac{105 - 110}{15/\sqrt{25}}\right) = P(Z > -\frac{5}{3}) \approx$

	.07
-1.6	.0475

$P(Z > -1.67) = 1 - .0475 = .9525$.

ex] Same as last ex but want $\alpha = 0.05$.

$.05 = \alpha = P_{100}(\bar{Y} > c) = P\left(\frac{\bar{Y} - 100}{15/\sqrt{25}} > \frac{c - 100}{15/\sqrt{25}}\right)$

$= P\left(Z > \frac{c - 100}{3}\right) = P(Z > z_{.95}) = P(Z > 1.645)$

	.04	.05
1.6	.9495	.9505

so $\frac{c - 100}{3} = 1.645$ or $c = 100 + 3(1.645) = 104.935$.

So $\beta(110) = P_{110}(\text{reject } H_0) = P_{110}(\bar{Y} > 104.935)$

$= P\left(\frac{\bar{Y}-110}{15/\sqrt{25}} > \frac{104.935-110}{15/\sqrt{25}}\right) \approx P(Z > -1.69)$

$= 1 - P(Z < -1.69) = 1 - .0455 = .9545$

α	$\beta(110)$
.0062	.7967
.0475	.9525
.05	.9545

trade off: as $\alpha \uparrow$ power $\uparrow$, as $\alpha \downarrow$ power $\downarrow$

§9.1 Not on final, sometimes on qual

1) $Y_1, ..., Y_n$ have joint pdf or pmf $f(\underset{\sim}{y}|\theta)$ for $\theta \in \Theta$. $[L_n, U_n]$ is a $100(1-\alpha)\%$ confidence interval CI for θ if $P_\theta(L_n(\underset{\sim}{Y}) \le \theta \le U_n(\underset{\sim}{Y})) = 1-\alpha$ for all $\theta \in \Theta$. $[L_n, U_n]$ is a large sample $100(1-\alpha)\%$ CI if $P_\theta(L_n(\underset{\sim}{Y}) \le \theta \le U_n(\underset{\sim}{Y})) \to 1-\alpha$ for all $\theta \in \Theta$.

2) $Y_1, ..., Y_n$ have joint pdf or pmf $f(\underset{\sim}{y}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta$. The random variable $R(\underset{\sim}{Y}|\underset{\sim}{\theta})$ is a pivot or pivotal quantity if the distribution of $R(\underset{\sim}{Y}|\underset{\sim}{\theta})$ is independent of $\underset{\sim}{\theta}$. $R(\underset{\sim}{Y}|\underset{\sim}{\theta})$ is an asymptotic pivot or asymptotical pivot quantity if the limiting distribution of $R(\underset{\sim}{Y}|\underset{\sim}{\theta})$ is independent of $\underset{\sim}{\theta}$. Then $R(\underset{\sim}{Y}|\underset{\sim}{\theta})$ is often useful for constructing a CI.

a statistic does not depend on θ

ex) $Y_1, \ldots, Y_n$ iid $N(\mu, \sigma^2)$

$$R(\underline{Y} \mid \mu, \sigma^2) = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t_{n-1} \xrightarrow{D} Z$$

$$1-\alpha = P\left(-t_{n-1, 1-\frac{\alpha}{2}} \le \frac{\bar{Y}-\mu}{S/\sqrt{n}} \le t_{n-1, 1-\frac{\alpha}{2}}\right)$$

$$= P\left(-t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \le \bar{Y} - \mu \le t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right)$$

$$= P\left(\bar{Y} - t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \le \mu \le \bar{Y} + t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right).$$

So $\bar{Y} \pm t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ is a $100(1-\alpha)\%$ CI for μ

where $P(t_{n-1} \le t_{n-1, 1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.

$\bar{Y} \pm z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ is a large sample $100(1-\alpha)\%$ CI for μ.

ex) See 9.1, 9.12.