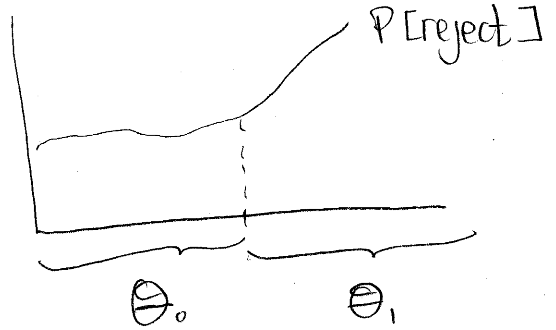


Occasionally, it may be the case that our test is optimal for any H_1.

Let $X_i \overset{iid}{\sim} N(\theta, 1)$. $H_0: \theta = \theta_0$, $H_1: \theta = \theta_1$.

By Neyman-Pearson, reject if $f(\bar{X}|\theta_1) > k f(\bar{X}|\theta_0)$

$$\Leftrightarrow \frac{f(\bar{X}|\theta_1)}{f(\bar{X}|\theta_0)} > k \Leftrightarrow \frac{(2\pi)^{-1/2}\left(\frac{1}{n}\right)^{-1/2} \exp\left\{-\frac{n}{2}(\bar{X}-\theta_1)^2\right\}}{(2\pi)^{-1/2}\left(\frac{1}{n}\right)^{-1/2} \exp\left\{-\frac{n}{2}(\bar{X}-\theta_0)^2\right\}} > k$$

$$\Leftrightarrow \exp\left\{-\frac{n}{2}\left[(\bar{X}-\theta_1)^2 - (\bar{X}-\theta_0)^2\right]\right\} > k$$

$$\Leftrightarrow \bar{X} > \frac{\frac{2\ln k}{n} + (\theta_1^2 - \theta_0^2)}{2(\theta_1 - \theta_0)} \quad \text{if } \theta_1 > \theta_0$$

$$\Leftrightarrow \underbrace{\sqrt{n}(\bar{X} - \theta_0)}_{\sim N(0,1) \text{ under the null}} > \underbrace{\sqrt{n}\left[\frac{\frac{2\ln k}{n} + (\theta_1^2 - \theta_0^2)}{2(\theta_1 - \theta_0)} - \theta_0\right]}_{1.645}$$

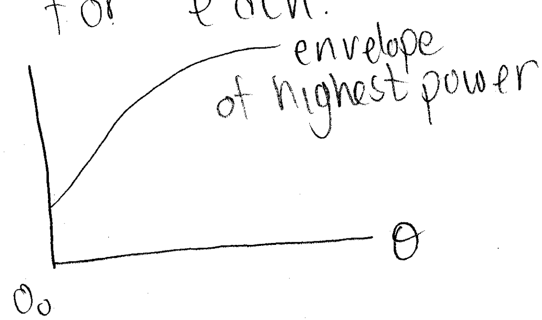
- use the same critical value for any $\theta_1 > \theta_0$.
- uniformly most powerful test.

The N-P test gives the same test $\forall \theta_1 > \theta_0$

If no uniformly most powerful test, still can use N-P Lemma.

1] Suppose $H_0: \theta = \theta_0$ vs $H_1: \theta > \theta_0$

Find N-P test for each $\theta_1 > \theta_0$. Compute $\beta(\theta)|_{\theta_1}$ for each.



◦ can use this as a benchmark

2] Point optimal test. Can find a test that is optimal somewhere.

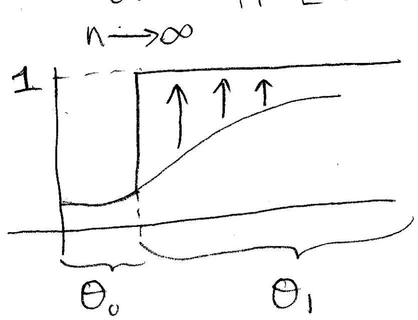
3] UMP unbiased test. Power $\geq$ size

Large samples

Defn: a hypothesis test testing $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$ is consistent if

$$\lim_{n\to\infty} \Pr[X \in R | \theta] = 1 \quad \forall \theta \in \Theta_1.$$

◦ You were right, Jason!



e.g. $X_i \overset{iid}{\sim} N(0,1)$. $H_0: \theta = \theta_0$, $H_1: \theta > \theta_0$

$$R = \{ \underline{X} ; \sqrt{n}\,(\bar{X} - \theta_0) > c \}$$

$$\Pr[\underline{X} \in R \mid \theta] = \Pr[\sqrt{n}\,(\bar{X} - \theta_0) > c \mid \theta]$$

$$= \Pr[\sqrt{n}\,(\bar{X} - \theta + \theta - \theta_0) > c \mid \theta]$$

$$= \Pr[\underbrace{\sqrt{n}\,(\bar{X} - \theta)}_{\sim N(0,1)} > \underbrace{c - \overbrace{\sqrt{n}\,(\theta - \theta_0)}^{\to +\infty}}_{\to -\infty} \mid \theta]$$

$$\to 1$$

Thus, this test is consistent.

Power against "local" alternative

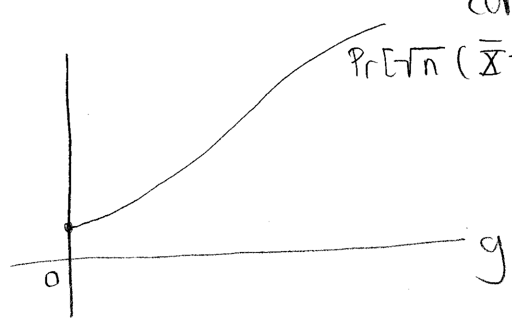
$H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_0 + g/\sqrt{n}$, so

$$\Pr[\sqrt{n}\,(\bar{X} - \theta) > c - \sqrt{n}\,(\theta - \theta_0) \mid \theta]$$

$$= \Pr[\sqrt{n}\,(\bar{X} - \theta) > c - \sqrt{n}\,(g/\sqrt{n}) \mid \theta]$$

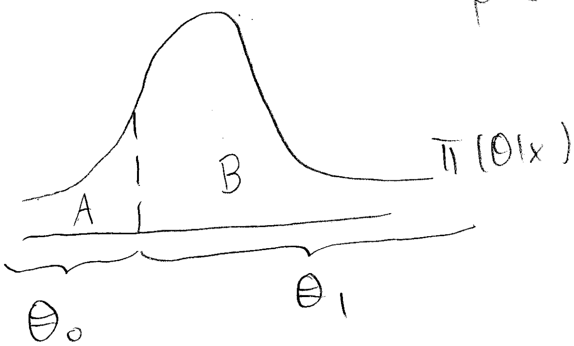
$$= \Pr[\sqrt{n}\,(\bar{X} - \theta) > \underbrace{c - g}_{\text{constant}} \mid \theta]$$

$\Pr[\sqrt{n}\,(\bar{X} - \theta) > c - g \mid \theta]$ = Local power function



Other approaches

<u>Bayesian</u>. Has prior $\pi(\theta)$
Sees data $f_X(x|\theta)$
Creates posterior $\pi(\theta|x) = \dfrac{f_X(x|\theta)\pi(\theta)}{\int f_X(x|\theta)\pi(\theta)d\theta}$



$$Pr[H_0 \text{ is true}] = \int_{\Theta_0} \pi(\theta|x)d\theta \equiv A$$

$$Pr[H_1 \text{ is true}] \equiv B$$

Decision rule: choose H_1 if $B > \frac{1}{k}A$

i.e. reject for $\dfrac{Pr[\theta \in \Theta_0]}{Pr[\theta \in \Theta_1]} < k$

<u>Remarks:</u>

1] This is referred to as the posterior odds ratio.

2] $k=1$ puts H_0, H_1 on equal footing.

<u>Large sample LR and equivalents</u>

- Need to assume a distribution, even if you don't believe it. (Quasi-MLE)

Thm: Let $X_1, \dots, X_n$ be a random sample with pdf $f(x|\theta)$. Consider test $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$. Under regularity conditions,

$$-2 \ln \lambda(x) \xrightarrow{d} \chi^2(1).$$

Remarks:

1) Heuristic of proof. Under some regularity conditions.

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_0} L(\theta|x)}{\sup_{\theta \in \Theta} L(\theta|x)} = \frac{L(\theta_0|x)}{L(\hat{\theta}_{MLE}|x)}$$

$$\ln \lambda(x) = \ln L(\theta_0|x) - \ln L(\hat{\theta}_{MLE}|x)$$

$$= \ln L(\hat{\theta}_{MLE}|x) + \underbrace{\frac{\partial}{\partial \theta} \ln L(\theta|x)\Big|_{\hat{\theta}_{MLE}}}_{=0} (\hat{\theta}_{MLE} - \theta_0)$$

$$+ \frac{\partial^2}{\partial \theta^2} \ln L(\theta|x)\Big|_{\hat{\theta}} (\hat{\theta}_{MLE} - \theta_0)^2 - \ln L(\hat{\theta}_{MLE}|x)$$

$$= \frac{1}{2} \frac{\partial^2}{\partial \theta^2} \ln L(\theta|x)\Big|_{\tilde{\theta}} (\hat{\theta}_{MLE} - \theta_0)^2$$

where $\hat{\theta} \in \text{conv}(\hat{\theta}_{MLE}, \theta_0)$

$$\Rightarrow -2 \ln \lambda(x) = -\frac{1}{n} \frac{\partial^2}{\partial \theta^2} \ln L(\theta|x)\Big|_{\hat{\theta}} \left[\sqrt{n}(\hat{\theta}_{MLE} - \theta_0)\right]^2$$

By CLT, $\sqrt{n}(\hat{\theta}_{MLE} - \theta_0) \xrightarrow{d} N(0, V(\theta))$ and

$$-\frac{1}{n} \frac{\partial^2}{\partial \theta^2} \ln L(\theta|x)\Big|_{\hat{\theta}} \xrightarrow{p} \text{constant}.$$

We know that $-\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\ln L(\theta|x)\Big|_{\tilde{\theta}} \to E\left[-\frac{\partial^2}{\partial\theta^2}\ln L_1(\theta_0|x)\right]$
$= I_1(\theta_0)$

Further, we know that $\sqrt{n}(\hat{\theta}_{MLE}-\theta_0) \xrightarrow{d} N(0,[I_1(\theta_0)]^{-1})$ assuming the conditions under which the MLE is asymptotically efficient.

Thus, $-2\ln\lambda(x) \xrightarrow{d} Y^2$, where $Y \sim N(0, I_1(\theta_0)\, I_1(\theta_0)^{-1})$
$\stackrel{d}{=} N(0,1)$

$\Rightarrow -2\ln\lambda(x) \xrightarrow{d} \chi^2(1)$

2] Regularity:
- 2-times differentiable $\log L(\cdot)$ (smoothness)
- $\hat{\theta}_{MLE}$ asymptotically efficient
- $\hat{\theta}_{MLE}$ asymptotically normal

Wald test

Recall: $-2\ln\lambda = -\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\ln L(\theta|x)\Big|_{\tilde{\theta}}\left(\sqrt{n}(\hat{\theta}_{MLE}-\theta_0)\right)^2$

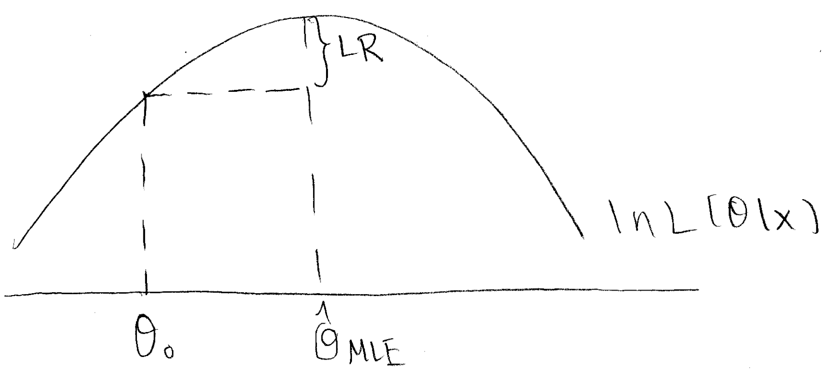
$$= \underbrace{-\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\ln L(\theta|x)\Big|_{\hat{\theta}}\left(\sqrt{n}(\hat{\theta}_{MLE}-\theta_0)\right)^2}_{\text{Wald test}} + o_p(1)$$

- we can actually compute this

Write $I_n(\theta) = E\left[-n \frac{\partial^2}{\partial \theta^2} \ln f(\bar{X} | \theta) \right]$

$$\hat{I}_n(\hat{\theta}_{ML}) = -n \frac{\partial^2}{\partial \theta^2} \ln f(\bar{X}|\theta) \Big|_{\hat{\theta}_{ML}}$$

$$\Rightarrow \quad W = n \hat{I}_1(\hat{\theta}_{ML}) (\hat{\theta}_{ML} - \theta_0)^2 \xrightarrow{d} \chi^2(1)$$



Wald takes $\hat{\theta}_{MLE} - \theta_0$ and transforms it using the curvature (Fisher information)