$$R(\theta, W(\cdot)) = E_\theta[L(\theta - W(\underline{X}))]$$
$$= \int [L(\theta - w(x))] f_{\underline{X}}(x|\theta)\,dx$$

Large sample criteria

a] Consistency

Defn 10.1.1: a sequence of estimators $\{W_n\}$ is consistent for θ if $\text{plim } W_n = \theta \ \forall \theta \in \Theta$.

Remarks:

1] as with $\xrightarrow{p}$ in general, this just captures the idea of "close."

2] Key difference: $\forall \theta \in \Theta$

3] Sense of closeness depends on method of convergence.

4] Recall from Chebyshev, if $\overbrace{E_\theta[(W_n - \theta)^2]}^{MSE} \to 0$ $\forall \theta \in \Theta$, then $\text{plim } W_n = \theta$

- $E_\theta[(W_n - \theta)^2] = Var_\theta(W_n) + [E[W_n] - \theta]^2$
- consistency is implied by $Var_\theta(W_n) \to 0$ and $Bias_\theta \to 0$.

5] Thm 10.1.5: If $\{W_n\}$ is consistent for θ and a_n is a sequence s.t. $a_n \to 1$ and $b_n \to 0$, then $U_n = a_n W_n + b_n$ is consistent for θ.

- Given one consistent estimator, there are a continuum of consistent estimators.

6] W_n consistent for θ,

$$U_n = \begin{cases} B_n & n \le 10{,}000 \\ W_n & n > 10{,}000 \end{cases}$$

U_n is also consistent for θ

7] Thm 10.1.6: MLEs under some regularity conditions are consistent.

$$\hat{\theta}_{MLE}(x) = \underset{\theta \in \Theta}{\operatorname{argmax}}\ L_n(\theta | x)$$

- Need Θ compact, L_n continuous in θ, L_n measurable in $\mathcal{X}$, and $L_n \to L$ uniformly in θ.

8] Method of moments: $\underbrace{m}_{\text{vector of moments}} = g(\hat{\theta}) \Rightarrow \hat{\theta} = g^{-1}(m)$

- if g^{-1} is nice enough, MM estimators are consistent.

eg. $X_i \overset{iid}{\sim}$ Bernoulli(p). Let $W_n(X) = \bar{X}_n$. Then

$$E[\bar{X}_n - p] = 0 \quad \forall p$$
$$Var((\bar{X}_n - p)) = \frac{p(1-p)}{n} \quad \forall p$$

Since $Var(\bar{X}_n) \to 0$, and bias $= 0$ $\forall n$, we have that $\operatorname{plim} \bar{X}_n = p$

5] Asymptotic Efficiency

Defn 10.1.1 A sequence of estimators $\{W_n\}$ is asymptotically efficient for $\tau(\theta)$ if

$$\sqrt{n}\,(W_n - \tau(\theta)) \xrightarrow{d} N(0, v(\theta)), \text{ where}$$

$$v(\theta) = \frac{(\tau'(\theta))^2}{E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right)^2\right]}$$

i.e. $\tau(\theta) = \theta \Rightarrow (\tau'(\theta))^2 = 1^2 = 1$

W_n has CLT and $\text{Var} = \frac{1}{I(\theta)}$ (CRLB attained asymptotically.)

e.g. $X_1, \ldots, X_n \overset{iid}{\sim}$ Bernoulli(p)

Let $W_n = \overline{X}_n$. $\text{Var}(\overline{X}_n) = \frac{p(1-p)}{n}$

$$\frac{\partial^2}{\partial p^2}\log L(p|x) = -\frac{n\overline{X}_n}{p^2} - \frac{n(1-\overline{X})}{(1-p)^2}$$

$$\Rightarrow E\left[-\frac{\partial^2}{\partial p^2}\log L(p|x)\right] = \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)} = I_n(p)$$

CLT: $\sqrt{n}\,(\overline{X}_n - p) \xrightarrow{d} N(0, p(1-p))$

Remarks:

1. Locally asymptotically normal.
2. iid Normal - typically get CRLB for all n.

Hypothesis Testing

Have a density $f(x|\theta)$. This is your economic model. $(\theta \in \Theta)$

a theory should restrict θ, ie $\theta \in \Theta_0 \subsetneq \Theta$.

<u>Defn</u>: a <u>statistical hypothesis</u> is a statement which implies that the true model belongs to a subset of possible models;

$$\Theta \supset \Theta_0 \cup \Theta_1 \text{ and } \Theta_0 \cap \Theta_1 = \emptyset.$$

$\theta \in \Theta_0$ = null hypothesis H_0

$\theta \in \Theta_1$ = alternative hypothesis H_1 or H_A

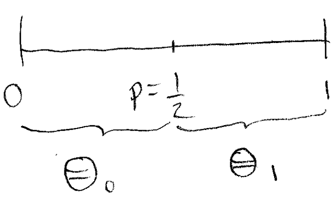
If $\Theta_0 = \{\theta_0\}$ (point), we refer to this as a point null, or simple null.

If $\Theta_1 = \{\theta_1\}$, we refer to this as a simple alternative.

<u>Defn</u>: a <u>hypothesis testing procedure</u> is a rule which specifies:

1. For which sample values we choose H_0
2. For which sample values we choose H_1.

e.g. $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$



$H_0: p \le \frac{1}{2}$

$H_1: p > \frac{1}{2}$

Say we use $\bar{X}_n$. "If $\bar{X}_n \le \frac{1}{2}$, choose H_0. If $\bar{X}_n > \frac{1}{2}$, choose H_1." This is a hypothesis testing procedure.

The sampling distribution of $\bar{X}_n$ leads to classification errors.

Rejection region: $\{(x_1, \dots, x_n) : \text{we choose } H_1\} \equiv R$

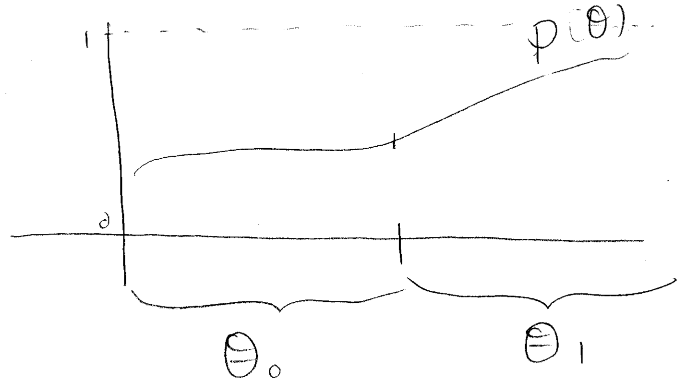
acceptance region: $\{(x_1, \dots, x_n) : \text{we choose } H_0\} \equiv A$

- We are going to make errors.

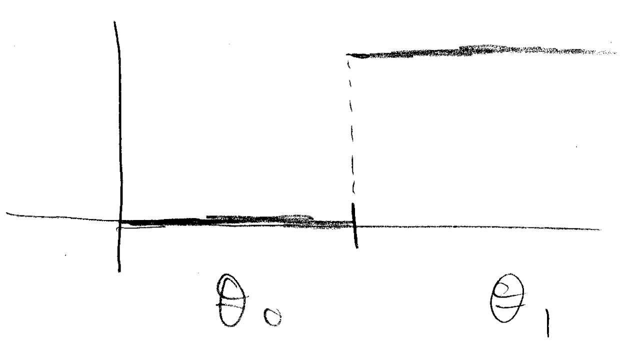
$\Pr[\text{Type 1 error}] = \Pr[\underline{X} \in R \mid \theta \in \Theta_0] \equiv \alpha(\theta)$

$\Pr[\text{Type 2 error}] = \Pr[\underline{X} \in A \mid \theta \in \Theta_1] \equiv 1 - \beta$

$\text{Power} = \Pr[\underline{X} \in R \mid \theta] \equiv \rho(\theta)$



Ideal test:



Classical approach: want to increase $p(\theta)$ $\forall \theta \in \Theta_1$ and decrease $p(\theta)$ $\forall \theta \in \Theta_0$

1] Fix "tolerable" level of type 1 error

- size
- level

2] Maximize power given size.