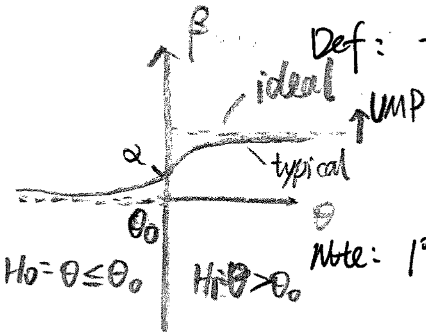
IV

Error Probabilities and Most Powerful Tests

Def: Type I Error means incorrectly rejecting H_0;

Type II Error means incorrectly accepting H_0.

Note: 1° With the above def., the prob. of Type I Error is $P_\theta(\underline{X} \in R)$, $\theta \in \Theta_0$; the prob. of Type II Error is $1 - P_\theta(\underline{X} \in R)$, $\theta \in \Theta_0^c$.



Def: The power function of a hypothesis test with rejection region R is

$$\beta(\theta) \equiv P_\theta(\underline{X} \in R)$$

Note: 1° The ideal power fn. is $\beta(\theta) = \begin{cases} 0, & \theta \in \Theta_0 \\ 1, & \theta \in \Theta_0^c \end{cases}$, which cannot be attained in most cases. A good test should have a power fn. as "close" to the ideal as possible.

2° Typically, the power fn. depends on the sample size n, so it can help determine an appropriate sample size.

3° For a fixed sample size, it's usually impossible to make both types of error prob. arbitrarily small. Naturally, we consider successive optimization.

Def: A test with power fn. $\beta(\theta)$ is a size α test, if $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$, $0 \le \alpha \le 1$.

Def: A test with power fn. $\beta(\theta)$ is a level α test, if $\sup_{\theta \in \Theta_0} \beta(\theta) \le \alpha$, $0 \le \alpha \le 1$.

Note: 1° Typical choices of α are 0.01, 0.05, 0.10.

2° For LRTs, a monotonic relation can be built between critical value c and size α. In general

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(\underline{X}) \leq c) = \alpha$$

3° Cutoff points: z_α — $P(Z > z_\alpha) = \alpha$, $Z \sim N(0,1)$

$t_{n-1,\frac{\alpha}{2}}$ — $P(T_{n-1} > t_{n-1,\frac{\alpha}{2}}) = \frac{\alpha}{2}$, $T_{n-1} \sim t_{n-1}$

$\chi^2_{p,1-\alpha}$ — $P(\chi^2_p > \chi^2_{p,1-\alpha}) = 1-\alpha$,

Def: A test with power fn. $\beta(\theta)$ is <u>unbiased</u> if $\beta(\theta') \geq \beta(\theta'')$, $\forall \theta' \in \Theta_0^c$, $\forall \theta'' \in \Theta_0$.

Def: Given hypotheses $H_0: \theta \in \Theta_0$; $H_1: \theta \in \Theta_0^c$,

C is a class of tests. A test in C with power fn. $\beta(\theta)$ is a <u>uniformly most powerful (UMP)</u> class C test, if

$$\beta(\theta) \geq \beta'(\theta), \ \forall \theta \in \Theta_0^c, \ \forall \beta'(\theta) \text{ of a test in } C.$$

Note: 1° UMP tests do not exist in many realistic problems.

2° A hypothesis is <u>simple</u>, if it corresponds to only one value of θ; alternatively, a <u>composite</u> hypothesis corresponds to more than one value of θ.

Thm: (Neyman-Pearson Lemma)

Given hypotheses $H_0: \theta = \theta_0$; $H_1: \theta = \theta_1$

and $\underline{X} \sim f(\underline{x} | \theta_i)$, $(i = 0,1)$.

Conditions:

(1) A test has rejection region and acceptance region:

$$R = \{\underline{x} \mid f(\underline{x}|\theta_1) > k f(\underline{x}|\theta_0)\}$$

$$R^c = \{\underline{x} \mid f(\underline{x}|\theta_1) < k f(\underline{x}|\theta_0)\}$$

where $k \geq 0$.

(2) $P_{\theta_0}(\underline{X} \in R) = \alpha$

Then:

a) (Sufficiency)

Any test that satisfies (1) and (2) is a UMP level α test.

b) (Necessity)

If there exists a test satisfying (1) and (2) with $k>0$, then every UMP level α test is a size α test, i.e. satisfies (2). and every UMP level α test satisfies (1) except for a measure-zero set on sample space, for both θ_0 and θ_1.

Note: 1° Condition (2) simply says the test is size α.

2° $R \cup R^c$ is not equal to the sample space, which is such stated for convenience of proof, but it's understandable.

Eg: 1° (UMP binomial test)

Given population $X \sim \text{Binomial}(2, \theta)$

Hypotheses are: $H_0: \theta = \frac{1}{2}$; $H_1: \theta = \frac{3}{4}$.

The ratios of pmfs are:

$$\frac{P(0|\theta=\frac{3}{4})}{P(0|\theta=\frac{1}{2})} = \frac{1}{4}, \quad \frac{P(1|\theta=\frac{3}{4})}{P(1|\theta=\frac{1}{2})} = \frac{3}{4}, \quad \frac{P(2|\theta=\frac{3}{4})}{P(2|\theta=\frac{1}{2})} = \frac{9}{4}$$

The test satisfying condition (1) in Neyman-Pearson Lemma with:

1° $\frac{3}{4} < k < \frac{9}{4}$, then $R = \{2\}$, $\alpha = P(2|\theta=\frac{1}{2}) = \frac{1}{4}$.

So the test is a UMP level $\frac{1}{4}$ test. (And the only one)

2° $\frac{1}{4} < k < \frac{3}{4}$, then $R = \{1, 2\}$, $\alpha = P(\{1,2\}|\theta=\frac{1}{2}) = \frac{3}{4}$.

So the test is a UMP level $\frac{3}{4}$ test. (And the only one

3° $0 < k < \frac{1}{4}$, then $R = \{0, 1, 2\}$, $\alpha = 1$.

4° $k > \frac{9}{4}$, then $R = \emptyset$, $\alpha = 0$.

2° (UMP normal test)

Given population $X \sim N(\theta, \sigma^2)$ with σ^2 known.

Hypotheses are: $H_0: \theta = \theta_0$; $H_1: \theta = \theta_1$ $(\theta_0 > \theta_1)$

We know $\bar{X}$ is a sufficient stat. for θ, and the ratio of pdfs is.

$$\frac{f_{\bar{X}}(\bar{x} \mid \theta = \theta_1)}{f_{\bar{X}}(\bar{x} \mid \theta = \theta_0)} = \exp\left\{ -\frac{n}{2\sigma^2}\left[2(\theta_0 - \theta_1)\bar{x} + (\theta_1^2 - \theta_0^2) \right] \right\}$$

The test based on $\bar{X}$ that satisfies condition (1) in Neyman-Pearson lemma gives rejection region:

$$S = \left\{ \bar{x} \;\middle|\; \bar{x} < \frac{-\sigma^2 \ln k}{n(\theta_0 - \theta_1)} + \frac{\theta_0 + \theta_1}{2} \right\} \quad (k > 0)$$

and it has size

$$\alpha = P_{\theta_0}(\bar{x} \in S) = \Phi\left(-\frac{\sigma \ln k}{\sqrt{n}(\theta_0 - \theta_1)} - \frac{\sqrt{n}(\theta_0 - \theta_1)}{2\sigma} \right)$$

By corollary of Neyman-Pearson lemma, this test is a UMP level α test.

Note: ~~1° Hypotheses that specify only one possible distn. for sample X are called simple hypotheses~~ ;

1° Hypotheses like $H: \theta \geq \theta_0$ or $H: \theta < \theta_0$ are called <u>one-sided hypotheses</u> ; Hypotheses that assert either a parameter is large or small, like $H: \theta \neq \theta_0$, are called <u>two-sided hypotheses</u>.

Def: r.v. $T \sim g(t \mid \theta)$, $\theta \in \Theta \subseteq \mathbb{R}$ has a <u>monotone likelihood ratio (MLR)</u>, if

$$\forall \theta_2 > \theta_1, \quad \frac{g(t \mid \theta_2)}{g(t \mid \theta_1)} \text{ is monotone on the union of supports of } g(t \mid \theta_2) \text{ and } g(t \mid \theta_1).$$

where $\frac{c}{0}$ is defined as $+\infty$ if $c > 0$.

Note: 1° ~~Normal~~ Gaussian (known variance), Poisson, and binomial have an MLR.

2° Any exponential family with $g(t|\theta) = h(t)\,c(\theta)\,e^{w(\theta)t}$ has an MLR if $w(\theta)$ is nondecreasing.

Thm: (Karlin - Rubin)

Suppose T is a sufficient stat. for θ, $T \sim g(t|\theta)$ has a nondecreasing MLR.

Hypotheses are $H_0: \theta \le \theta_0$, $H_1: \theta > \theta_0$

Then $\forall\, t_0$, the test with rejection region $R = \{\underline{x} : T(\underline{x}) > t_0\}$ is a UMP level $\alpha = P_{\theta_0}(T > t_0)$ test.

Eg: 1° (UMP normal test)

Population $X \sim N(\theta, \sigma^2)$, σ^2 known, $\Theta = \mathbb{R}$.

Hypotheses are $H_0: \theta \ge \theta_0$; $H_1: \theta < \theta_0$.

Since $\overline{X}$ is sufficient for θ, and has an MLR, the Karlin-Rubin thm give that the test with

$$R = \{\underline{x} : \bar{x} < -\frac{\sigma z_\alpha}{\sqrt{n}} + \theta_0\}$$

is a UMP level α test.