NP Lemma helps us find the Most Powerful test

for $H_0: \theta = \theta_0$

vs $H_1: \theta = \theta_1$

Uniformly Most Powerful test.

$H_0: \theta = \theta_0$

vs $H_1: \theta > \theta_0$

[$H_0: \theta = \theta_0$

vs $H_1: \theta \neq \theta_0$]

test function $\phi(\underset{\sim}{X}) = \begin{cases} 1 & \underset{\sim}{X} \in RR \\ 0 & \underset{\sim}{X} \notin RR \end{cases}$

1. control prob. of Type I error (size) : $\alpha \geq E_{\theta_0}[\phi(\underset{\sim}{X})] = P_{\theta_0}(\underset{\sim}{X} \in RR)$

2. for any other test function: $\phi^*(\underset{\sim}{X}) = \begin{cases} 1 & \underset{\sim}{X} \in RR^* \\ 0 & \underset{\sim}{X} \notin RR^* \end{cases}$

with $E_{\theta_0}[\phi^*(\underset{\sim}{X})] \leq \alpha$.

If type II error of $\phi(\underset{\sim}{X})$ is smaller than the one of $\phi^*(\underset{\sim}{X})$

for any $\theta = \theta_1 > \theta_0$

$$\begin{cases} P_{\theta_1}(X \notin RR) = 1 - P_{\theta_1}(X \in RR) = 1 - E_{\theta_1}(\phi(\underset{\sim}{X})) \\ \leq P_{\theta_1}(X \notin RR^*) = 1 - E_{\theta_1}(\phi^*(\underset{\sim}{X})) \end{cases}$$

$$E_{\theta_1}(\phi(\underset{\sim}{X})) \geq E_{\theta_1}(\phi^*(\underset{\sim}{X}))$$

- For simple vs one-sided alternative. to find UMP can be started from finding the MP test for simple vs simple.

Two-sided case

$H_0: \theta = \theta_0$

vs $H_1: \theta \neq \theta_0$.

start with Simple vs Simple pick a value θ_1 in $\{\theta: \theta \neq \theta_0\}$

$H_0: \theta = \theta_0$
vs $H_1: \theta = \theta_1 > \theta_1$

UMP
for $H_0: \theta = \theta_0$
vs $H_1: \theta > \theta_0$

$H_0: \theta = \theta_0$
vs $H_1: \theta = \theta_1 < \theta_0$

✓

UMP
for $H_0: \theta = \theta_0$
vs $H_1: \theta < \theta_0$

UMP for $H_0: \theta \leq \theta_0$
vs $H_1: \theta > \theta_0$

If you want to use NP Lemma,

pick $\theta_0^* \leq \theta_0$ pick $\theta_1^* > \theta_0$

$H_0: \theta = \theta_0^*$
vs $H_1: \theta = \theta_1^*$

} fails

- MLR

$g_T(t|\theta)$

for any $\theta_2 > \theta_1$ $\dfrac{g_T(t|\theta_2)}{g_T(t|\theta_1)} \nearrow t$

8.3.3 Uniformly most powerful tests

Remark: So far, we have discussed "test related optimality" in the context of simple-versus-simple hypotheses. We now extend the idea of "most powerful" to more realistic situations involving composite hypotheses; e.g., $H_0: \theta \leq \theta_0$ versus $H_1: \theta > \theta_0$.

Definition: A family of pdfs (pmfs) $\{g_T(t|\theta); \theta \in \Theta\}$ for a univariate random variable T has **monotone likelihood ratio (MLR)** if for all $\theta_2 > \theta_1$, the ratio

$$\frac{g_T(t|\theta_2)}{g_T(t|\theta_1)}$$

is a nondecreasing function of t over the set $\{t : g_T(t|\theta_1) > 0 \text{ or } g_T(t|\theta_2) > 0\}$.

Example 8.15. Suppose $T \sim b(n, \theta)$, where $0 < \theta < 1$. The pmf of T is

$$g_T(t|\theta) = \binom{n}{t}\theta^t(1-\theta)^{n-t},$$

for $t = 0, 1, 2, ..., n$. Suppose $\theta_2 > \theta_1$. Consider

$$\frac{g_T(t|\theta_2)}{g_T(t|\theta_1)} = \frac{\binom{n}{t}\theta_2^t(1-\theta_2)^{n-t}}{\binom{n}{t}\theta_1^t(1-\theta_1)^{n-t}} = \left(\frac{1-\theta_2}{1-\theta_1}\right)^n \left[\frac{\theta_2(1-\theta_1)}{\theta_1(1-\theta_2)}\right]^t.$$

Note that $\left(\frac{1-\theta_2}{1-\theta_1}\right)^n > 0$ and is free of t. Also, because $\theta_2 > \theta_1$, both

$$\frac{\theta_2}{\theta_1} > 1 \quad \text{and} \quad \frac{1-\theta_1}{1-\theta_2} > 1.$$

Therefore,

$$\frac{g_T(t|\theta_2)}{g_T(t|\theta_1)} = \underbrace{c(\theta_1, \theta_2)}_{>0} a^t,$$

where $a > 1$. This is an increasing function of t over $\{t : t = 0, 1, 2, ..., n\}$. Therefore, the family $\{g_T(t|\theta) : 0 < \theta < 1\}$ has MLR.

Remark: Many common families of pdfs (pmfs) have MLR. For example, if

$$T \sim g_T(t|\theta) = h(t)c(\theta)e^{w(\theta)t},$$

i.e., T has pdf (pmf) in the one-parameter exponential family, then $\{g_T(t|\theta); \theta \in \Theta\}$ has MLR if $w(\theta)$ is a nondecreasing function of θ.
Proof. Exercise.

Q: Why is MLR useful?
A: It makes getting UMP tests easy.

Theorem 8.3.17 (Karlin-Rubin). Consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

Suppose that T is sufficient. Suppose that $\{g_T(t|\theta); \theta \in \Theta\}$ has MLR. The test that rejects H_0 iff $T > t_0$ is a UMP level α test, where

$$\alpha = P_{\theta_0}(T > t_0).$$

Similarly, when testing

$$H_0 : \theta \geq \theta_0 \quad \text{versus} \quad H_1 : \theta < \theta_0,$$

the test that rejects H_0 iff $T < t_0$ is UMP level α, where $\alpha = P_{\theta_0}(T < t_0)$.

Example 8.16. Suppose $X_1, X_2, ..., X_n$ are iid Bernoulli(θ), where $0 < \theta < 1$, and consider testing

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

We know that

$$T = \sum_{i=1}^{n} X_i$$

is a sufficient statistic and $T \sim b(n, \theta)$. In Example 8.15, we showed that the family $\{g_T(t|\theta) : 0 < \theta < 1\}$ has MLR. Therefore, the Karlin-Rubin Theorem says that the UMP level α test is

$$\phi(t) = I(t > t_0),$$

where t_0 solves

$$\alpha = P_{\theta_0}(T > t_0) = \sum_{t=\lfloor t_0 \rfloor + 1}^{n} \binom{n}{t} \theta_0^t (1 - \theta_0)^{n-t}.$$

Special case: I took $n = 30$ and $\theta_0 = 0.2$. I used R to calculate the following:

t_0	$P_{\theta_0}(T \geq \lfloor t_0 \rfloor + 1)$
$7 \leq t_0 < 8$	$P(T \geq 8 \mid \theta = 0.2) = 0.2392$
$8 \leq t_0 < 9$	$P(T \geq 9 \mid \theta = 0.2) = 0.1287$
$9 \leq t_0 < 10$	$P(T \geq 10 \mid \theta = 0.2) = 0.0611$
$10 \leq t_0 < 11$	$P(T \geq 11 \mid \theta = 0.2) = 0.0256$
$11 \leq t_0 < 12$	$P(T \geq 12 \mid \theta = 0.2) = 0.0095$

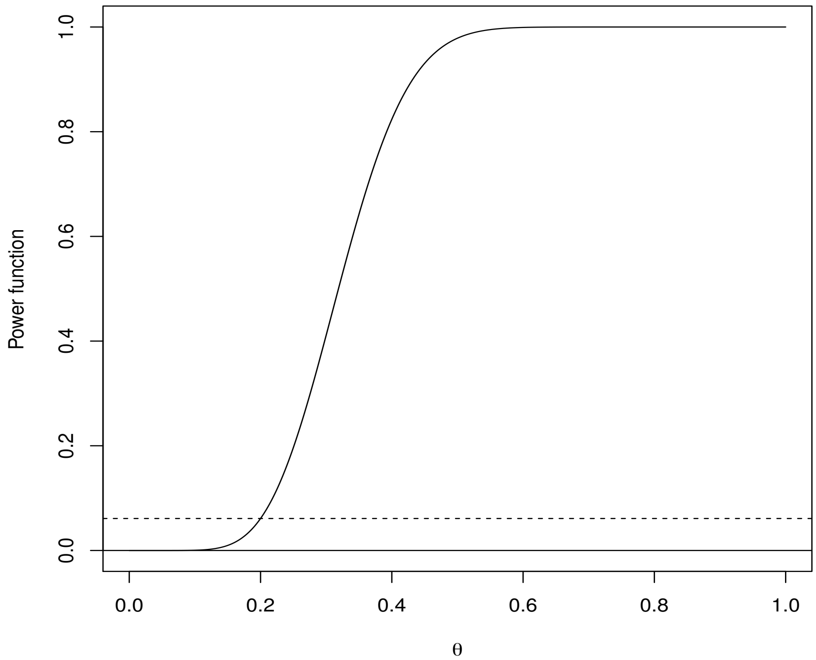

Figure 8.4: Power function $\beta(\theta)$ for the UMP level $\alpha = 0.0611$ test in Example 8.16 with $n = 30$ and $\theta_0 = 0.2$. A horizontal line at $\alpha = 0.0611$ has been added.

Therefore, the UMP level $\alpha = 0.0611$ test of $H_0 : \theta \leq 0.2$ versus $H_1 : \theta > 0.2$ uses $I(t \geq 10)$. The UMP level $\alpha = 0.0256$ test uses $I(t \geq 11)$. Note that (without randomizing) it is not possible to write a UMP level $\alpha = 0.05$ test in this problem. For the level $\alpha = 0.0611$ test, the power function is

$$\beta(\theta) = P_\theta(T \geq 10) = \sum_{t=10}^{30} \binom{30}{t} \theta^t (1-\theta)^{30-t},$$

which is depicted in Figure 8.4 (above).

Example 8.17. Suppose that $X_1, X_2, ..., X_n$ are iid with population distribution

$$f_X(x|\theta) = \theta e^{-\theta x} I(x > 0),$$

where $\theta > 0$. Note that this population distribution is an exponential distribution with mean $1/\theta$. Derive the UMP level α test for

$$H_0 : \theta \geq \theta_0$$
versus
$$H_1 : \theta < \theta_0.$$

Solution. It is easy to show that

$$T = \sum_{i=1}^{n} X_i$$

is a sufficient statistic and $T \sim \text{gamma}(n, 1/\theta)$. Suppose $\theta_2 > \theta_1$ and form the ratio

$$\frac{g_T(t|\theta_2)}{g_T(t|\theta_1)} = \frac{\dfrac{1}{\Gamma(n)\left(\frac{1}{\theta_2}\right)^n} t^{n-1} e^{-\theta_2 t}}{\dfrac{1}{\Gamma(n)\left(\frac{1}{\theta_1}\right)^n} t^{n-1} e^{-\theta_1 t}} = \left(\frac{\theta_2}{\theta_1}\right)^n e^{-t(\theta_2-\theta_1)}.$$

Because $\theta_2 - \theta_1 > 0$, we see that the ratio

$$\frac{g_T(t|\theta_2)}{g_T(t|\theta_1)}$$

is a decreasing function of t over $\{t : t > 0\}$. However, the ratio is an increasing function of $t^* = -t$, and $T^* = T^*(\mathbf{X}) = -\sum_{i=1}^{n} X_i$ is still a sufficient statistic (it is a one-to-one function of T). Therefore, we can apply the Karlin-Rubin Theorem using $T^* = -T$ instead. Specifically, the UMP level α test is

$$\phi(t^*) = I(t^* < t_0),$$

where t_0 satisfies

$$\begin{aligned} \alpha = E_{\theta_0}[\phi(T^*)] &= P_{\theta_0}(T^* < t_0) \\ &= P_{\theta_0}(T > -t_0). \end{aligned}$$

Because $T \sim \text{gamma}(n, 1/\theta)$, we take $-t_0 = g_{n,1/\theta_0,\alpha}$, the (upper) α quantile of a $\text{gamma}(n, 1/\theta_0)$ distribution. Therefore, the UMP level α test is $I(t > g_{n,1/\theta_0,\alpha})$; i.e., the UMP level α rejection region is

$$R = \left\{\mathbf{x} \in \mathcal{X} : \sum_{i=1}^{n} x_i > g_{n,1/\theta_0,\alpha}\right\}.$$

Using χ^2 critical values: We can also write this rejection region in terms of a χ^2 quantile. To see why, note that when $\theta = \theta_0$, the quantity $2\theta_0 T \sim \chi^2_{2n}$ so that

$$\begin{aligned} \alpha = P_{\theta_0}(T > -t_0) &= P_{\theta_0}(2\theta_0 T > -2\theta_0 t_0) \\ &\Longrightarrow -2\theta_0 t_0 \stackrel{\text{set}}{=} \chi^2_{2n,\alpha}. \end{aligned}$$

Therefore, the UMP level α rejection region can be written as

$$R = \left\{\mathbf{x} \in \mathcal{X} : 2\theta_0 \sum_{i=1}^{n} x_i > \chi^2_{2n,\alpha}\right\} = \left\{\mathbf{x} \in \mathcal{X} : \sum_{i=1}^{n} x_i > \frac{\chi^2_{2n,\alpha}}{2\theta_0}\right\}.$$

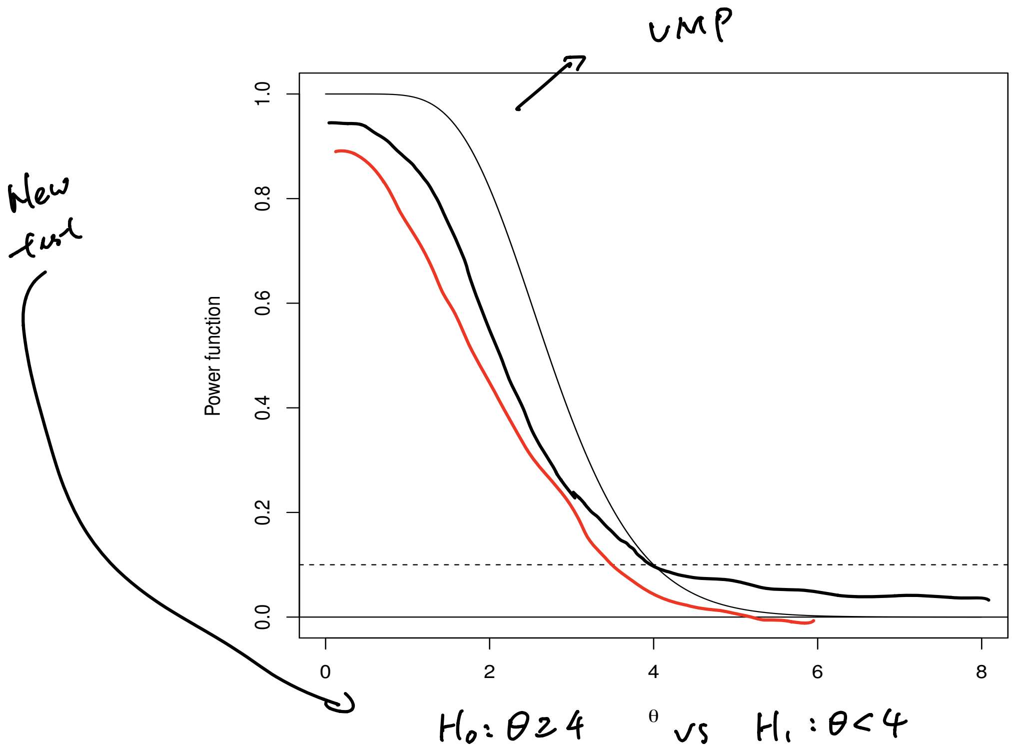

Figure 8.5: Power function $\beta(\theta)$ for the UMP level $\alpha = 0.10$ test in Example 8.17 with $n = 10$ and $\theta_0 = 4$. A horizontal line at $\alpha = 0.10$ has been added.

Remark: One advantage of writing the rejection region in this way is that it depends on a χ^2 quantile, which, historically, may have been available in probability tables (i.e., in times before computers and R). Another small advantage is that we can express the power function $\beta(\theta)$ in terms of a χ^2 cdf instead of a more general gamma cdf.

Power function: The power function of the UMP level α test is given by

$$\begin{aligned}\beta(\theta) = P_\theta(\mathbf{X} \in R) = P_\theta\left(T > \frac{\chi^2_{2n,\alpha}}{2\theta_0}\right) &= P_\theta\left(2\theta T > \frac{\theta\chi^2_{2n,\alpha}}{\theta_0}\right) \\ &= 1 - F_{\chi^2_{2n}}\left(\frac{\theta\chi^2_{2n,\alpha}}{\theta_0}\right),\end{aligned}$$

where $F_{\chi^2_{2n}}(\cdot)$ is the χ^2_{2n} cdf. A graph of this power function, when $n = 10$, $\alpha = 0.10$, and $\theta_0 = 4$, is shown in Figure 8.5 (above).

Proof of Karlin-Rubin Theorem. We will prove this theorem in parts. The first part is a lemma.

Lemma 1: If $g(x) \uparrow_{\text{nd}} x$ and $h(x) \uparrow_{\text{nd}} x$, then

$$\text{cov}[g(X), h(X)] \geq 0.$$