UNIFORMLY MOST POWERFUL (UMP) TESTS [1]

ECONOMICS G6411
Class Notes (November 2000; revised July 2004)

PHOEBUS J. DHRYMES

This is to **supplement** not to **supplant** the discussion in *Gallant.* . . . .

We always operate in the following context: All random variables are defined on the probability space $(\Omega , \ \mathcal{A} , \ \mathcal{P})$, are continuous and are endowed with density functions, denoted generically by f . We are dealing with a random sample, i.e. a sequence of independent, identically distributed (i.i.d.) random variables whose density is characterized by a scalar parameter denoted generically by θ which is an element of the admissible parameter space $\Theta \subseteq R$.

Definition 1. A statistical hypothesis is a (usually numerical) statement regarding the parameter θ of the p.d.f. f . The normal context is to express the admissible space $\Theta = \omega \cup \bar{\omega}$, where $\bar{\omega}$ is the complement of ω **in** Θ , and to state the hypothesis as

$H_0 : \theta \in \omega$
as against
$H_1 : \theta \in \bar{\omega}$.

The hypothesis of interest H_0 is said to be the **null hypothesis**, while H_1 is said to be the **alternative hypothesis**. Finally if ω consists of a single point, the null hypothesis is said to be **simple**, otherwise it is said to be **composite**; similarly for the alternative hypothesis.

We now particularize the context, so that given a sequence of random variables $X_1, \ldots, X_n$ we consider the probability space induced by the sequence, $(R^n, \mathcal{B}(R^n), P_\theta)$, where P_θ is a family of distribution functions indexed by the parameter θ .

Definition 2. A **test of a hypothesis** is a procedure which enables us to accept or reject the null hypothesis once a sample is obtained. More specifically, let $S_r \subset R^n$, $\bar{S}_r = S_a$ its complement in R^n and let $x = (x_1, x_2, \ldots, x_n)'$ be a sample (observation). The sets S_r , S_a are said to be, respectively, the **rejection** and **acceptance** region. If $x \in S_r$ the null hypothesis is **rejected**, while if $x \in S_a$ the null hypothesis is **accepted**.

Notice that once the rejection region is defined, the **test is defined**; notice further

---

[1]
The structure of this exposition follows the discussion in George G. Roussas *A Course in Mathematical Statistics*, Academic Press: San Diego, 1997, pp. 337-346.

that the probability of rejecting the null hypothesis is simply

$$\pi_{S_r}(\theta) = \int_{S_r} dP_\theta .$$

We have

Definition 3. The function $\pi_{S_r}(\cdot)$ is said to be the[2] **power function** corresponding to the test defined by the rejection region S_r, and always gives the probability of rejection, given that the true parameter is the one specified by its argument. If the argument is the value specified by the null hypothesis, the function yields the probability of rejecting a **true** null, or the probability of type I error, also termed the **size** or **level of significance** of the test. If it is evaluated at an argument specified by the alternative hypothesis, the function yields the probability of rejecting a **false null**, or the **power of the test** relative to this alternative; finally, one minus this quantity is said to be the probability of a type II error, or the probability of accepting a **false null**.

As an example, suppose $\omega = \theta_0$, and $\bar{\omega} = \theta_1$, so that both the null and the alternative are **simple** hypotheses. Let S_r be the rejection region of this test. Then

$$\pi_{S_r}(\theta_0), \quad \pi_{S_r}(\theta_1), \quad 1 - \pi_{S_r}(\theta_1),$$

are, respectively, the level of significance (probability of type I error), the power of the test relative to the alternative θ_1, and the probability of type II error (acceptance of a false null).

A typical (non-randomized) test involves the definition of a function (of the sample)

$$\begin{aligned} \phi(x) &= 1 \quad \text{if} \quad x \in S_r \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{1}$$

It is then apparent that

$$\pi_{S_r}(\theta) = \int_{R^n} \phi(x) d\, P_\theta(x) = E_\theta[\phi] \tag{2}$$

Definition 4. An α level test defined by the rejection region S_r is said to be **uniformly most powerful** for testing the null $\theta \in \omega$, as against $\theta \in \bar{\omega}$ if and only if

i. $\sup_{\theta \in \omega} \pi_{S_r}(\theta) = \alpha$;

ii. for any other test (of level α or less) defined by the rejection region S_r^*

$$\pi_{S_r}(\theta) \geq \pi_{S_r^*}(\theta), \quad \text{for all} \quad \theta \in \bar{\omega}.$$

---

[2] Often the term power function is restricted to the evaluation of $\pi_{S_r}(\theta)$, for $\theta \in \bar{\omega}$)

If $\omega = \{\theta_0\}$, $\bar{\omega} = \{\theta_1\}$, such a test is said to be **most powerful**.

A routine way of determining (most powerful) tests is through the following

Proposition 1. (Neyman-Pearson Lemma) Let $X = (X_1, X_2, \ldots, X_n)'$ be a sequence of i.i.d. random variables with density $f(\cdot;\theta)$ $\theta \in \Theta$, sometimes also denoted by f_θ . Let the joint density of the observations (the likelihood function) be denoted by

$$L(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta), \quad \text{or} \quad L_\theta, \tag{3}$$

and suppose we are interested in testing the **simple** null $\theta = \theta_0$, as against the **simple** alternative $\theta = \theta_1$, at the level of significance $\alpha \in (0,1)$. For some constants $\lambda > 0$, $c > 0$ define the function

$$\begin{aligned} \phi(x) &= 1, \quad \text{if} \quad L(x;\theta_1) > \lambda L(x;\theta_0) \\ &= c, \quad \text{if} \quad L(x;\theta_1) = \lambda L(x;\theta_0) \\ &= 0, \quad \text{otherwise}, \end{aligned} \tag{4}$$

where λ and c are to be determined by the condition

$$E_{\theta_0}\phi(X) = \int_{S_r} \phi(x) d\, P_{\theta_0}(x) = \alpha, \tag{5}$$

where S_r is the rejection region defined by the test ϕ , and is given by

$$S_r = \{x : L(x;\theta_1 > \lambda L(x:\theta_0)\} \cup \{x : L(x:\theta_1) = \lambda L(x:\theta_0)\}.$$

The test defined by (the rejection region) S_r is the most powerful (MP) α-level test for testing the null θ_0 as against the alternative $\theta = \theta_1$, in the sense that for any other test, ϕ^* which defines (another) rejection region S_r^*

$$\pi_{S_r}(\theta_1) \geq \pi_{S_r^*}(\theta_1). \tag{6}$$

Proof: Let $L(x;\theta_0)$, $L(x;\theta_1)$ denote the likelihood function of the observations[3] under the null θ_0 , and alternative θ_1 , and for ease of notation write them as L_0, L_1 , respectively. Given footnote 3, we can then write

$$\begin{aligned} E_{\theta_0}\phi(X) &= \Pr\left(\frac{L_1}{L_0} > \lambda|\theta_0\right) + c\Pr\left(\frac{L_1}{L_0} = \lambda|\theta_0\right) \\ &= P_{\theta_0}(Y > \lambda) + cP_{\theta_0}(Y = \lambda), \quad Y = \frac{L_1}{L_0}. \end{aligned} \tag{7}$$

---

[3] Because at some stage we will have to divide by $L(x;\theta_0)$, we need to make sure that this entity is not zero. To this end, we may assume the set of points in R^n , for which this occurs has P_{θ_0} -measure zero, and redefine the domain and/or the function accordingly, so that the operation above is permissible.

If we put

$$J(\lambda) = P_{\theta_0}(Y > \lambda) \text{ and } G(\lambda) = 1 - J(\lambda), \tag{8}$$

then $G(\cdot)$ is the **distribution function** of Y, and $J(\cdot)$ is a non-decreasing right continuous function such that $J(-\infty) = 1$ and $J(\infty) = 0$. Moreover,

$$P_{\theta_0}(Y = \lambda) = G(\lambda) - G(\lambda^-) = J(\lambda^-) - J(\lambda), \tag{9}$$

where λ^- is the left limit; note further that by construction, for any $0 < \alpha < 1$, there exists $\lambda_0 \geq 0$, such that

$$J(\lambda_0^-) \leq \alpha \leq J(\lambda_0). \tag{10}$$

If λ_0 is a **continuity** point of J, then $\alpha = J(\lambda_0)$, and since the test is of size α, and we may take [4]

$$c = 0. \tag{11}$$

Next, suppose that λ_0 above is a point of **discontinuity**; take

$$c = \frac{\alpha - J(\lambda_0)}{J(\lambda_0^-) - J(\lambda_0)} \tag{12}$$

so that

$$E_{\theta_0}(\phi) = J(\lambda_0) + \frac{\alpha - J(\lambda_0)}{J(\lambda_0^-) - J(\lambda_0)}[J(\lambda_0^-) - J(\lambda_0)] = \alpha, \tag{13}$$

so we have again a level α test. This completes the construction of the (level α) test because we have already implicitly defined the rejection region as

$$S_r = \left(x : \frac{L(x;\theta_1)}{L(x;\theta_0)} \geq \lambda. \right) \tag{14}$$

It remains now to show that this test is most powerful within the class of tests whose level is equal to or less than α. Let ϕ^* define another test of level equal to or less than α and let S_r^* be its rejection region. Let

$$B^+ = \{x : \phi(x) - \phi^*(x) > 0\}, \quad B^- = \{x : \phi(x) - \phi^*(x) < 0\}$$

$$B^0 = \{x : \phi(x) - \phi^*(x) = 0, \} \tag{15}$$

and note that

$$R^n = B^+ \cup B^- \cup B^0, \tag{16}$$

is a **partition** of R^n, i.e. the sets in the right member of Eq. (16) are **disjoint** and, moreover

$$B^+ \subseteq \{x : L_1(x) \geq \lambda L_0(x)\}, \quad B^- \subseteq \{x : L_1(x) \leq \lambda L_0(x)\}. \tag{17}$$

Consequently,

$$\int_{R^n} [\phi(x) - \phi^*(x)][L_1(x) - \lambda L_0(x)]dx = \int_{B^+} [\phi(x) - \phi^*(x)][L_1(x) - \lambda L_0(x)]dx \tag{18}$$

$$+ \int_{B^-} [\phi(x) - \phi^*(x)][L_1(x) - \lambda L_0(x)]dx \geq 0,$$

---

[4] Notice that if the function $J(\cdot)$ is continuous in its argument, then Eq. (4) may be modified so that $\phi(x) = 1$ if $L_1/L_0 \geq \lambda$, and zero otherwise.

because the integral over B^0 is **zero** and the integrands of the two integrals of the right member of Eq. (18) are **non-negative**, so that both integrals in the right member are non-negative. On the other hand

$$\int_{R^n} [\phi(x) - \phi^*(x)] L_0(x) dx = \pi_{S_r}(\theta_0) - \pi_{S_r^*}(\theta_0) \geq 0, \tag{19}$$

because the test with rejection region S_r is of **size** α , while the test with rejection regions S_r^* (ϕ^*) is of **size equal to or less than** α . Next consider

$$\int_{R^n} [\phi(x) - \phi^*(x)] L_1(x) dx = \pi_{S_r}(\theta_1) - \pi_{S_r^*}(\theta_1) \geq 0, \tag{20}$$

because, from Eq. (18),

$$\int_{R^n} [\phi(x) - \phi^*(x)] L_1(x) dx \geq \lambda \int_{R^n} [\phi(x) - \phi^*(x)] L_0(x) dx \geq 0. \tag{21}$$

Corollary 1. For the test above with rejection region S_r $\pi_{S_r}(\theta_1) \geq \alpha$, i.e. the power of this test against the alternative is **at least** α .

Proof: Consider the trivial test

$$\phi^*(x) = \alpha,$$

which has a rejection region, say S_r^* . Since $\pi_{S_r^*}(\theta_1) = \alpha$, and the test with rejection region S_r is **most powserful**, we must thus have

$$\pi_{S_r}(\theta_1) \geq \pi_{S_r^*}(\theta_1) = \alpha. \tag{22}$$

q.e.d.

0.1 Uniformly Most Powerful (UMP) Tests

We begin with a definition.

Definition 2. Let X be a family of i.i.d. random variables, with density $f(\cdot;\theta)$ and likelihood function $L(\cdot;\theta)$. It is said to have the **monotone likelihood ratio** (MLR) property in V if and only if there exists a function V **independent of parameters** such that, for $\theta_1, \theta_2 \in \Theta$ the associated likelihood ratio is a monotone function of V .

Example 1. Consider an i.i.d. sequence of normal variables with mean θ and variance one, $X = (X_1, X_2, \ldots, X_n)$. Its likelihood function is given by

$$L(x;\theta) = [2\pi]^{-(n/2)} \prod_{i=1}^{n} e^{-(1/2)(x_i-\theta)^2} = [2\pi]^{-(n/2)} e^{-(1/2)\sum_{i=1}^{n} x_i^2} \times e^{-(n/2)\theta^2} \times e^{-(n/2)\theta\bar{x}}$$

$$= k_1(x) k_2(\theta) e^{-(n/2)\theta\bar{x}}. \tag{23}$$

The ratio of these likelihood functions evaluated at $\theta_1 < \theta_2$ is given by

$$\frac{L_1}{L_2} = \left(\frac{k_2(\theta_1)}{k_2(\theta_1)}\right) e^{-(n/2)(\theta_1-\theta_2)\bar{x}}. \tag{24}$$

Since $\theta_1 < \theta_2$ the ratio is **monotone increasing** in $V(x) = \bar{x}$. If $\theta_1 > \theta_2$, the ratio would be **monotone decreasing**. Thus, variables with normal distribution (and known variance) have the MLR property, as is the case with normal variables with known mean and unknown variance.

In fact it may be shown that variables having a density of the form (the so called exponential family)

$$f(x;\theta) = h(x)K(\theta)e^{Q(\theta)T(x)} \tag{25}$$

have likelihood ratio functions with the MLR property.

We may now deal with the construction of UMP tests.

We have

Proposition 2. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with density $f(\cdot;\theta)$, $\theta \in \Theta \subseteq R$, that has the MLR property. Let $\theta_0 \in \Theta$ and define $\omega = \{\theta : \theta \in \Theta, \text{ and } \theta \leq \theta_0, \}$, $\bar{\omega} = \{\theta : \theta \in \Theta, \text{ and } \theta > \theta_0, \}$. For testing the hypothesis

$H_0 : \theta = \theta_0$
as against the alternative
$H_1 : \theta \in \bar{\omega}, (\theta > \theta_0)$

define the test

$$\begin{aligned} \phi(x) &= 1, \quad \text{if } V(x) > \lambda \\ &= c \quad \text{if } V(x) = \lambda \\ &= 0, \quad \text{otherwise}, \end{aligned} \tag{26}$$

where λ, c are uniquely defined by

$$\pi_{S_r}(\theta_0) = E_{\theta_0}\phi(X) = P_{\theta_0}(V(x) > \lambda) + cP_{\theta_0}(V(x) = \lambda) = \alpha \tag{27}$$

The α-level test above, (defined by its rejection region $S_r = \{x : V(x) \geq \lambda\}$) is most powerful (MP) within the class of tests of level equal to less than α for testing the simple hypothesis H_0 against the composite alternative H_1.

Proof: Let $\theta_1 \in \bar{\omega}$ be arbitrary, and consider the simple null $\theta = \theta_0$ as against the simple alternative $\theta = \theta_1$. By Proposition 1 the MP α-level test is given by

$$\begin{aligned} \phi_1(x) &= 1, \quad \text{if } \frac{L_1}{L_0} > \lambda_1^* \\ &= c_1^* \quad \text{if } \frac{L_1}{L_0} = \lambda_1^* \\ &= 0, \quad \text{otherwise}, \end{aligned} \tag{28}$$

such that

$$\pi_{S_r^*}(\theta_0) = E_{\theta_0}\phi_1(X) = P_{\theta 0}\left(\frac{L_1}{L_0} > \lambda_1^*\right) + cP_{\theta_0}\left(\frac{L_1}{L_2} = \lambda_1^*\right) = \alpha. \tag{29}$$

Eq. (29) will determine uniquely λ_1^* and c_1^*. To use the MLR property first put

$$\nu(V(x)) = \frac{L_1}{L_2} \tag{30}$$

and notice that because ν is **monotoone** in V and thus invertible, the test defined by the conditions

$$\nu(V(x)) > \lambda_1^*, \quad \nu(V(x)) = \lambda_1^* \tag{31}$$

may also be rendered as

$$S_r^* = \{x : \nu(V(x)) \geq \lambda_1^*\} = \{x : V(x) \geq \nu^{-1}(\lambda_1^*) = \lambda_1\}. \tag{32}$$

Thus, Eq. (29) may also be written as

$$\pi_{S_r^*}(\theta_0) = P_{\theta_0}(V(x) > \lambda_1) + cP_{\theta_0}(V(x) = \lambda_1) = \alpha, \tag{33}$$

and λ_1, c are uniquely determined by Eq. (33). Comparing Eq. (33) with Eq. (27) it is evident that $\lambda_1 = \lambda, c_1 = c$. It then follows immediately that the test in Eq. (28) is **independent** of θ_1; since the latter is an arbitrary element of $\bar{\omega}$ it follows that the test in Eq. (28) is the same as the test in Eq. (26), which is, thus, the UMP test relative to alternatives in $\bar{\omega}$.

q.e.d.

Remark 2. It may be shown that for the class of densities having the MLR property, the power function is non-decreasing in θ, i.e.

$$\pi_{S_r}(\theta_1) \geq \pi_{S_r}(\theta_2), \quad \text{if } \theta_1 \geq \theta_2, \tag{34}$$

a fact illustrated by the Proposition 3, below.

Proposition 3. Under the conditions stated in Proposition 2, the test defined by Eqs. (26) and (27) has the property that

$$E_{\theta^*}\phi(X) = \pi_{S_r}(\theta^*) = P_{\theta^*}(V(x) > \lambda) + cP_{\theta^*}(V(x) = \lambda) \leq \alpha, \quad \text{for } \theta^* \in \omega. \tag{35}$$

Proof: Let ϕ^* be the test defined by

$$\begin{aligned} \phi^* = 1, &\quad \text{if } \nu(V(x)) = \frac{L_0}{L_1} > \lambda^{**} \\ = c^{**} &\quad \text{if } \nu(V(x)) = \lambda^{**} \\ = 0, &\quad \text{othersie.} \end{aligned} \tag{36}$$

Because of the MLR property ν is invertible, and the function in Eq. (36) may also be stated as

$$\begin{aligned}\phi^* &= 1, \quad \text{if } \ V(x) > \lambda^* = \psi^{-1}(\lambda^{**}) \\ &= c^{**} \quad \text{if } \ V(x) = \lambda^* \\ &= 0, \quad \text{otherwise.}\end{aligned} \tag{37}$$

The test above defines the rejection region $S_r^* = \{x : V(x) \geq \lambda^*\}$, and the entities λ^*, c^* are uniquely determined by the condition

$$\pi_{S_r^*}(\theta^*) = E_{\theta^*}\phi^*(X) = P_{\theta^*}(V(x) > \lambda^*) + P_{\theta^*}(V(x) = \lambda^*) = \alpha(\theta^*). \tag{38}$$

If we evaluate Eq. (27) at θ^* we obtain

$$\pi_{S_r}(\theta^*) = P_{\theta^*}(V(x) > \lambda) + cP_{\theta^*}(V(x) = \lambda) = \alpha(\theta^*). \tag{39}$$

Comparing Eqs. (39) and (38), we conclude that that $\lambda = \lambda^*$, $c = c^*$, and the two tests are identical. Moreover, by Corollary 1, we conclude that

$$\alpha(\theta^*) \leq \alpha, \quad \text{since } \ \alpha = \alpha(\theta_0), \tag{40}$$

and the latter is the power of the test defined in this proposition.

q.e.d.

Finally, we have

Proposition 4. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with density $f(\cdot;\theta)$, $\theta \in \Theta \subseteq R$, that has the MLR property. Let $\theta_0 \in \Theta$ and define $\omega = \{\theta : \theta \in \Theta, \text{ and } \ \theta \leq \theta_0,\}$, $\bar{\omega} = \{\theta : \theta \in \Theta, \text{ and } \ \theta > \theta_0,\}$. For testing

$H_0 : \theta \in \omega$
as against the alternative
$H_1 : \theta \in \bar{\omega}$

there exists, within the class of all tests with level of significance $\leq \alpha$, a **UMP** test defined by the function

$$\begin{aligned}\phi(x) &= 1, \quad \text{if } \ V(x) > \lambda \\ &= c, \quad \text{if } \ V(x) = \lambda \\ &= \quad \text{otherwise,}\end{aligned} \tag{41}$$

where λ, c are (uniquely) determined by the condition

$$E_{\theta_0}\phi(X) = P_{\theta_0}(V(x) > \lambda) + cP_{\theta_0}(V(x) = \lambda) = \alpha, \tag{42}$$

provided the likelihood ratio is **increasing** in V. If it is **decreasing**, the inequalities above are reversed.

Proof: The proof follows immediately from Propositions 2 and 3. Thus, let

$$\mathcal{C} = \text{the class of } \alpha\text{-level tests of the hypothesis } \ H_0 : \theta \in \omega;$$

$$\mathcal{C}_0 = \text{the class of } \alpha\text{-level tests of the hypothesis } \ H_0 : \theta = \theta_0. \tag{43}$$

If a test, $\phi \in \mathcal{C}$, then it certainly belongs in $\mathcal{C}_0$; consequently $\mathcal{C} \subseteq \mathcal{C}_0$. The test defined by Eqs. (26) and (27), (or Eqs. (41) and (42), belongs in $\mathcal{C}$ by Proposition 3, and by Proposition 2 it is uniformly most powerful in the class $\mathcal{C}_0$. Hence within the class of tests of level $\leq \alpha$, the test defined by Eqs. (26) and (27) is UMP.

q.e.d.

Example 2. Consider again the situation in Example 1, where $V(x) = \bar{x}$, and $\bar{x} \sim N(\theta, \frac{1}{n})$. The test defined in Proposition 2 for $\alpha = .05$ and $n = 100$, is given by the rejection region determined by

$$P_{\theta_0}(\bar{x} \geq \lambda) = .05. \tag{44}$$

For $\theta_0 = 0$, the rejection region is

$$S_r = \{x : \sqrt{n}\bar{x} \geq 1.96\} = \{x : \bar{x} \geq .196\}. \tag{45}$$

To illustrate Proposition 3, we need to show that this rejection region (test) has a **level** $\leq \alpha$ for testing the null hypothesis $\theta = \theta^* < 0$. The level of significance of this test (as a test of the hypothesis just noted) is

$$P_{\theta^*}(\bar{x} \geq .164) \leq .05. \tag{46}$$

This is so **because** under the null hypothesis $\theta = \theta^* < 0$

$$\bar{x} - \theta^* \sim N(0, \frac{1}{n}),$$

so that

$$P_{\theta^*}\left(\sqrt{n}(\bar{x} - \theta^*) \geq 1.64\right) = P_{\theta^*}(\bar{x} \geq .164 + \theta^*) = \alpha = .05. \tag{47}$$

But for the test defined by the critical region $S_r\{x : \bar{x} \geq .164\}$ we must have

$$P_{\theta^*}(\bar{x} \geq .164) \leq P_{\theta^*}(\bar{x} \geq .164 + \theta^*) = \alpha = .05, \tag{48}$$

because $\theta^* < \theta_0 = 0$, which demonstrates the meaning of Proposition 3.